2003 Annual Proceedings - Anaheim: Volume #1

of
Selected Research and Development Papers
Presented at
The National Convention of the Association for Educational Communications and Technology
Sponsored by the Research and Theory Division

Anaheim, CA
2003

Editors
Michael Simonson
Program Professor
Instructional Technology and Distance Education
Nova Southeastern University
Fischler Graduate School of Education and Human Services
North Miami Beach, FL
and
Margaret Crawford
Information Specialist
Mason City Public Schools
Mason City, IA

Previous Proceedings Published in ERIC

Year	Location	ED Number
1979	New Orleans	171329
1980	Denver	194061
1981	Philadelphia	207487
1982	Dallas	223191 – 223326
1983	New Orleans	231337
1984	Dallas	243411
1985	Anaheim	256301
1986	Las Vegas	267753
1987	Atlanta	285518
1988	New Orleans	295621
1989	Dallas	308805
1990	Anaheim	323912
1991	Orlando	334969
1992	Washington, D.C.	347970 – 348041
1993	New Orleans	362144
1994	Nashville	373774
1995	Anaheim	383284
1996	Indianapolis	397772
1997	Albuquerque	409832
1998	St. Louis	423819
1999	Houston	436128
2000	Long Beach	444595
2000	Denver	455756
2001	Atlanta	470066
2002	Dallas	

Preface

For the twenty-sixth year, the Research and Theory Division of the Association for Educational Communications and Technology (AECT) is sponsoring the publication of these Proceedings. Papers published in this volume were presented at the National AECT Convention in Anaheim, CA. A limited quantity of these Proceedings were printed and sold in both hardcopy and electronic versions.copies of both volumes were distributed to Convention attendees on compact disk. volume #1 is also available on microfiche through the Educational Resources Clearinghouse (ERIC) System.

The Proceedings of AECT's Convention are published in two volumes. Volume #1 contains papers dealing primarily with research and development topics. Papers dealing with instruction and training issues are contained in volume #2 which also contains over 60 papers.

REFEREEING PROCESS: Papers selected for presentation at the AECT Convention and included in these Proceedings were subjected to a reviewing process. All references to authorship were removed from proposals before they were submitted to referees for review. Approximately fifty percent of the manuscripts submitted for consideration were selected for presentation at the convention and for publication in these Proceedings. The papers contained in this document represent some of the most current thinking in educational communications and technology.

M. R. Simonson
Editor

2003 AECT Conference RTD Reviewer

Wilhelmina Savenye
Rhonda Robinson
Ward Cates
Susan Land
Michele Domisch
Kathy Wagner
Bob Hannafin
Harriette Spiegel
Charlie Reigeluth
Frank Harvey
Rhonda Robinson
Steve Crooks
Jonathan Brinkerhoff
Shujen Chang
Walter Dick
Kathy Schuh
Amy Leh
Bernda Bannan-Ritland
Kathryn Wong
John Burton
Barbara Lockee
Kathy Cennamo
Glen Holmes

Nancye McCrary
Chris Miller
Hossein Vaez
Gary Anglin
Doug Smith
Joan Mazur
Andrea Peach
Nathan Lowell
David Jonassen
Mike Hannafin
Al Januszewski
Deborah Lowther
Greg Sherman
Mike Simonson
Scott Adams
Ali Carr
Gary Morrison
Robert Reiser
Melissa Gibson
Stephen Gance
Carol Koroghlanian
Marcy Driscoll
Denis Hlynka

Brock Allen
Edward Caffarella
Ann De Vaney
Philip Doughty
Frank Dwyer
Peg Ertmer
Scott Grabinger
Janette R. Hill
Simon Hooper
P.K. Jamison
Frank Harvey
Rita Richey
Barbara Seels
Marty Tessmer
Brent WIlson
Tim Newby
Tom Reeves
Heidi Schnackenberg
Delia Neuman
Loyd Rieber
Xiadong Lin
William Win
Kevin Moore
Richard Howell

Order Form

Name:____

Affiliation:______________________________________

Address:_______________________________________

City, State, Zip: ___________________________________

Shipping Address:__________________________________

Shipping City, State, Zip: ______________________________

Email Address: ____________________________________

Phone Number: ____________________________________

Additional Phone Number:______________________________

Orders:

Please fill in and return, with payment, to Michael Simonson, Nova Southeastern University
1750 NE 167th Street. North Miami Beach, FL 33162
Make checks payable to 'Proceedings'.
VOLUME #1: RESEARCH PAPERS

Printed Version: Number of Copies_____________ @$80.00 each, Total __________

Electronic Version: Number of Copies_____________ @$80.00 each, Total __________

Both the Printed and Electronic: @ 120.00 Combined , Total__________

VOLUME #2: INSTRUCTION PAPERS
Electronic Version (No Printed Version): Number of Copies_____________ @$80.00 each, Total __________

Total Enclosed: $__________

NOTE: Previous Proceedings are available, also.

Table of Contents

Successful Development Strategies in Online Teacher Education Workshops

Jacquelyn Abromitis
Christine Clarke
University of Northern Iowa

Abstract

The Department of Teaching at University of Northern Iowa (UNI) offers two online programs for teachers: Online Professional Development Workshops (Prof Dev) and Professional Teacher Competencies (PTC) Workshops. The Prof Dev workshops are aimed at practicing professionals who are seeking continuing professional development for personal or professional reasons. The PTC Workshops were developed to encompass the UNI Professional Teacher Competencies, which identify 14 categories of skills that professional teachers should exhibit. The workshops are primarily used in a Masters in Learning and Technology (MLT) program at Western Governors University (WGU).

The Prof Dev Program began with one workshop in the Summer of 1997. Of the 36 students in the workshop, only three did not complete the workshop. The workshop was offered for a second time in Spring 1998. The attrition rate was approximately the same. Enthusiasm for the workshop coupled with low attrition led to three additional workshops. In Summer of 1998, four different workshops were offered. In Summer 2003, 21 workshops will be offered through this program. Throughout the rapid growth of the Prof Dev program, a completion with passing grade rate of 95% has been achieved, ranging from 92 – 96%.

The PTC workshops began in May 2002. As an Educational Provider to Western Governors University, UNI provides seats available in the UNI PTC workshops to WGU students in the MLT program. While seven workshops were written for the PTC program, at first three were selected as appropriate matches for the WGU program. After several months a fourth workshop was added to the workshop offerings. In summer 2002 (May – August start and complete dates), WGU placed 40 students in the PTC workshops. In Spring 2003 (workshops beginning and ending January – May), 95 students from WGU were placed into PTC workshops. Of the 95, five withdrew while the workshop was in session, five had "Incomplete" as a final grade, and two had "F" as a final grade. Eighty-three students completed the workshop with passing grades.

Workshop Development Strategies

Several workshop development strategies have contributed to the success of both the Prof Dev and PTC workshops: uniformity of design within workshops, uniformity of design between workshops, independence of workshop delivery tools, development of learning communities, paced asynchronous design, constructivist principles, and a faculty mentoring program.

Uniformity of Design Within Workshops

A key element to assist students in acclimating quickly to the workshop Web pages is the uniform design of content categories for each module of instruction within each workshop. For example, in Learning Styles: Instructional Strategies Using Technologies, there are four modules of instruction: Theories of Learning, Learning Styles, Instructional Strategies, and Evaluation. Within each module the following topics are covered: introduction, readings, videos, activities, and evaluation. Therefore, a simple navigation structure is maintained on every workshop Web page. Also available on every page are links to home, workshop info, the workshop calendar, and a site map. Workshop info gives the students quick links to the instructor's email address, the email address for the workshop, and the workshop roster (including student email addresses and short bios).

Uniformity of Design Between Workshops

The navigation and design of the workshop Web pages are maintained from workshop to workshop. In the PTC workshops, for example, a consistent color scheme is used with different banners for each workshop. The number of modules varies from workshop to workshop, but the pages within each workshop remain consistent. The consistency of design is very comfortable for the students, very few of whom report "lost in cyberspace" issues.

Independence of Workshop Delivery Tools

UNI does not use any workshop delivery tools such as WebCT or Blackboard. By creating Web pages in HTML, UNI is free to include only what is needed on the workshop Web pages and can include features potentially not available in packaged workshop delivery tools. By creating pages independently, UNI is able to create aesthetically pleasing Web pages rather than being limited to “canned” designs. Web pages can be customized based on best practices of Web page development, such as limiting the horizontal length of a line of text. Additionally, budget issues and university contracts with vendors are of no concern. Switching vendors would require moving all workshop data from one vendor’s product to another, resulting in tremendous re-work for the workshop developers and additional learning curves for the students. By being independent of vendor products, these issues are avoided.

Development of Learning Communities

The development of learning communities is an extremely important element of UNI’s workshop development strategies. Initially, students complete an introductory activity so that everyone in the workshop has basic information on workshop participants (location, grade levels taught, extra activities, and personal interests). Also, collaborative activities are integrated extensively throughout each workshop. Collaboration is a key element in helping students facilitate each other’s learning. Peer review teams are established by instructors based on students’ common interests or subject matter taught. This further enhances professional collaboration and learning. Also, each workshop instructor facilitates discussions based on student input to current topics. By doing so, the instructor is able to create collaborative webs of information based on individual student input. This, too, helps alleviate feelings of isolation.

Paced Asynchronous Design

Another important workshop development strategy is that each workshop is designed to be paced but asynchronous. In other words, each workshop has beginning and ending dates for the workshop and beginning and ending dates for each assignment; however, students may work at their own pace within those constructs. Students are not required to be online at any specified time. Asynchronous paced instruction allows students freedom from working at specific times, but also gives them deadlines to help them stay on schedule. Individual needs are addressed by instructors, such as making allowances for students who fall behind or need to work ahead.

Faculty spend a great deal of energy working with students who do not keep pace with the workshop. In order to keep students on schedule and motivated, faculty email students every several days if they have not heard from them. If the faculty member still does not hear from the student, he/she phones the student. Additionally, the director of the program contacts the student’s WGU mentor to have the mentor also make contact with the student. This is an extensive, time-consuming process, but it is one of the most effective aspects of keeping students on schedule and keeping attrition rates low.

Each faculty member in both programs has experience as a student in an online workshop. UNI believes that the best online faculty are those who first were online students. This sensitizes faculty to the needs of online learners.

Constructivist Principles

All workshops in both the Prof Dev and PTC programs are based on constructivist principles. All participants are practicing teachers, many with 15 to 20 years of classroom experience. Therefore, using constructivist principles of building upon previous knowledge was a natural fit in the philosophy of development. In each workshop, teachers create a project to be used in their classrooms. Often, they adapt or modify existing classroom materials.

Faculty Mentoring Program

Mentoring occurs between faculty members. New faculty will often observe or co-teach a workshop with an experienced faculty member. A faculty-only mailing list allows faculty to share ideas and concerns related not only to the workshop subject matter but also issues related to online learning. The mailing list functions as ongoing professional development for the faculty. Additionally, faculty also communicate privately with each other by chat and by phone.

The director of the program also mentors faculty. A faculty resources Web page is maintained to help faculty find relevant support materials. For example, a checklist of critical junctures in a workshop is available to faculty on the faculty resources page. Instructions for using the workshop electronic mailing list and support for using the online gradebook are also provided. Model introductory letters and workshop materials are continually being updated and supplemented by the director of the program.

Faculty workshop evaluations are provided at the completion of each workshop. Student names and email addresses are removed to maintain anonymity. Evaluations are reviewed between the director of the program and the faculty member.

Evaluation of Workshops and Faculty

Both a midpoint and a final workshop evaluation are administered in every workshop. The midpoint evaluation is intended to be formative, while the end of workshop evaluation is intended to be summative. The final workshop evaluation in particular is lengthy and consists of not only multiple choice questions but also 14 short answer questions. The feedback from these evaluations has contributed extensively to the continued revisions of the workshops. Based on student responses, the most important element of the UNI online workshops is quick, quality feedback from instructors.

Conclusion

The workshop development strategies outlined above have contributed not only to steady, quick growth, but also to an impressively low attrition rate in both the Prof Dev and PTC programs. Student testimonials can be viewed on the introductory page for each workshop. Workshop introductory pages for the PTC program can be found at http://www.uni.edu/teaching/testimonials.html. Student testimonials for the Prof Dev program can be found at http://www.uni.edu/profdev/gallery/comments.html.

Additional information can be found regarding the Prof Dev workshops at http://www.uni.edu/profdev/ and the PTC workshops at http://www.uni.edu/teaching/.

Cognitive Apprenticeship in Online Computer Training: What are Effective Facilitative Techniques for Online Workshops?

Paul Alford
Semiral Öncü
Kyong-Jee Kim
Indiana University

Abstract

This paper presents a pilot study attempting to measure the learning gain, satisfaction, and transfer of knowledge in online computer training workshops being conducted during the summer of 2003 in the IT Training and Education program at Indiana University. One control and one treatment student groups took an online workshop over a two week period. Teaching with cognitive apprenticeship was emulated in both groups with exploration activities and strategies utilized in one group and de-emphasized in the control group. Findings did not support the hypothesis that exploration strategies and techniques would result in increased learning gains and satisfaction, though increased interaction was observed and may lead to other implications for building an online program.

Introduction

What makes for good design in online training? Specifically, how should online training for Information Technology workshops be structured, implemented, and facilitated? This is the driving question of focus for this research study. The impetus for this research is the recent desire to take residential IT workshops offered by IT Training and Education at Indiana University and offer them in an online facilitated fashion.

The IT Training and Education program within University Information Technology Services at IU has been offering computer workshops to the IU community for over 15 years. These short, hands-on, instructor-led skill-based workshops provide training and learning experiences for specific computer applications and skills.

In order to extend these workshop offerings to a larger audience, provide more convenient training opportunities to existing clients, and enhance the learning environment to provide more and richer learning opportunities, it has been proposed that many of these workshops be offered online in a Distance Education mode. This would provide asynchronous and potentially some synchronous learning with an instructor to provide direction and mentoring, learning with the materials, and interaction with other students.

However, the materials used in this program are designed for instructor-led, hands-on workshop delivery in a computer classroom. This design does not translate well to online learning. In addition to the design and development of online materials, the facilitation of these online workshops should also support the unique environment provided by online interactions. This is also quite different from the residential environment and the same model employed in the classroom will not be appropriate for online facilitation.

Theoretical Framework

In order to provide an adequate delivery model for these workshops, several models were considered and reviewed. The most promising model identified was that of cognitive apprenticeship, first proposed by Collins, Brown and Newman (1989). Cognitive Apprenticeship was put forth as an update to traditional apprenticeship, specifically emphasizing training in the cognitive aspects of most activities, and as an alternative to the current view of information and skills being abstracted from their traditional context and being artificially reconstructed in classroom settings. Several have argued that, by returning this learning to appropriate social contexts, learners will better grasp both the content and the application potential of this knowledge (Brown et. al. 1989). The idea of an apprenticeship model, particularly emphasizing cognitive tasks, coincides well with IT Training where specific skills and applications are often taught for the purpose of increasing both skills and productivity for common tasks. As students learn the basics, they may also benefit from seeing expert practice modeled and explained in context by advanced users (Collins 1991).

Sociocultural theory

The roots of cognitive apprenticeship are in the social cognition theories of Vygotsky and others (Brown et. al. 1989). In his sociocultural theory, Vygotsky (1979) defines learning as a process of problem solving. A problem can be solved either individually or with the help of someone else. The relation between the learner and the helper can then be represented as the relation between a master and an apprentice. Using these terms to represent learners' positions, Vygotsky denotes this relation in three specific concepts (Mitchell and Myles, 1998). The first concept is *regulating the problem*. A problem is regulated either by *self regulation* where the learner is not in need of someone else's assistance, or by *other regulation* where the learner is assisted as we discussed above. The second concept is scaffolding. In scaffolding, the learner is supported by the master in several different ways such as notifying the gap between the current and the ideal versions of a performance, and making things easier. Third, one of the better-known concepts associated with Vygotsky is the Zone of Proximal Development (ZPD), defined as "the distance between a child's independent problem-solving level and that obtained under adult guidance or in collaboration with more capable peers (Bonk, 1998, p. 37)." It is evident that Vygotsky's sociocultural theory conceives how learners construct knowledge through interaction with other people that forms the culture.

Situated learning

Situated learning has developed based on Gibson's theory of affordances, Schoenfeld's theory on mathematical problem solving, and Vygotsky's social learning (Lave, 1988). It is also linked with Luira (1979) and Leont'ev (1981) (in Hendricks, 2001, p. 302). In situated learning, the instruction is related to what learners need in the real life situations (Stein, 1982). There are four basic principles in situated learning (Lave, 1988): (1) learning takes place in situations that we encounter in real life situations, (2) knowledge is situation specific and can only be transferred to similar conditions, (3) learning is a product of social process, and (4) learning exists in social environments, which is also based on Vygotsky's notion that learning and doing are cultural processes that cannot be separated from each other. Stein (1982) considers these four elements as "opportunities for educators to engage adult learners in … selecting situations that will engage learners in complex and realistic problem-centered activities … and [in] fostering the notion of cognitive apprenticeships" (p. 1). Lave (1988) studied the situated learning in the context of "community of practice" and pointed out that social interactions are key components of situated learning. Brown, Collins, and Duguid (1989) later developed the idea of situated learning into the cognitive apprenticeship framework.

Cognitive apprenticeship

In connection with situated learning and based on the idea that "learning is most effective when it approximates real-world situations or problem scenarios (Bonk, 1998, p.38)", cognitive apprenticeship emerged as an alternative to traditional teaching, learning and apprenticeship. Pioneering work on the notion and scope of apprenticeship was observed and conducted by Jean Lave (Brown et. al. 1989). Several others extended these principles, though not under the banner of 'cognitive apprenticeship', including Annemarie Palincsar's research on reciprocal teaching which provided evidence for the effectiveness of role-reversal (Palincsar, 1984). Collins, Brown, and Newman (1989) proposed cognitive apprenticeship as a new model for teaching and learning. The basic premise behind cognitive apprenticeship is that teachers should have their students as apprentices where teachers work alongside students and prepare situations in which learners engage in activities even if they cannot understand the basic conception behind them. Teachers become models for the students and students are provided with support based on the methods associated with cognitive apprenticeship. Collins et al. (1989) and Collins, Brown, and Holum (1991) outline cognitive apprenticeship in four dimensions: content, method, sequence, and sociological aspects; but cognitive apprenticeship is well-known especially with the methods of modeling, coaching, articulation, reflection, and exploration. Modeling has to do with an expert's carrying out a task for students. The task is usually an example of what students are about to perform. In educational settings typically teachers are the experts to perform such tasks. Performing the job externalizes the cognitive processes that are otherwise hidden and makes it possible for students to observe. Coaching involves monitoring students while they are performing the task, and offering any kind of support including modeling, scaffolding and so forth to improve their performances nearer to the expert performance. Scaffolding is any help either physically or verbally to support students while they are carrying out the task. Scaffolding may call for teachers to perform most of the carrying out as the students are not able to do it; however this should be faded out as the students advance. Articulation entails promoting students to enunciate their takings, way of thinking, and their solutions, basically their internal cognitive processes. The purpose of articulation is to have students enhance their understandings by articulating individually or by explaining to others as experts do in their modeling performances. Reflection refers to the act that enables students to compare their performances with each other and the expert. Various techniques can be appointed for replaying the think aloud procedures engaged by the students, which results in reflection. And finally, exploration involves encouraging students to problem-solve on their own. In order to provide this, teachers define the general goals and students are promoted to give attention to the particular parts of those goals. This also is to help students to employ newly learned skills on different examples.

Several researchers have discussed the use of cognitive apprenticeship as a technique in different contexts. Wilson, Jonassen, and Cole (1993) consider cognitive apprenticeship one of the various models of instruction that help transfer skills and knowledge to job environments. de Haan (2002) compares cognitive apprenticeship with his model of "knowledge, learning, and guidance" and reports that distributed notion of cognition, as promoted in cognitive apprenticeship, is a promising model for teaching. Wilson et al. (1993) compares cognitive apprenticeship as an instructional design along with various other instructional design models and contends that cognitive apprenticeship has both advantageous and disadvantageous properties. Collins (1991) argues that current technologies make learning possible thorough apprenticeship techniques more feasible and cost effective than before.

While cognitive apprenticeship was first suggested as an alternative model to classroom educational practices, it was quickly adopted as a good potential model for automated and online learning environments (Collins 1991, Teles 1993). Early uses of cognitive apprenticeship give several success examples either by design or coincidence (Berryman 1991, Brown et. al 1989). On the other hand, several recent studies report results from cognitive apprenticeship that are mostly empirical. Johnson and Fischbach (1993), for instance, studied coaching, modeling and fading in association with cognitive apprenticeship in college industrial technology classes and found that students taught with cognitive apprenticeship did slightly better than the traditionally taught students on problem solving and final exam; conversely, the traditional teaching methods were slightly better in standardized test settings – results were not statistically significant. Teles (1993) provides examples of successful cognitive apprenticeship techniques, especially scaffolding, reflection, and exploration through mentorship and peer collaboration in online learning and collaborative environments. Elliott (1994) case-studied rapid decision making in cognitive apprenticeship settings and identified kinds of comments an effective teacher made, the teacher's sources of information, and the patterns of the decision making during a particular reading lesson. Hendricks (2001) found evidence to indicate cognitive apprenticeship can be an effective teaching and learning tool, compared to abstracted learning, for short-term learning gains; though, this study did not show any difference between cognitive apprenticeship techniques and abstracted learning techniques for long-term memory and transfer of knowledge. Casey's (1996) findings indicate that multimedia applications of cognitive apprenticeship can support needs of diverse learners. Especially, students' tendency to refer to the expert was of importance of the findings. Derry and Lesgold (1996) inform that report that retention of knowledge learned with cognitive apprenticeship was perfect – less than 10 percent of information was lost after a six month retention test. They also report that 20-25 hours of training with cognitive apprenticeship is worth four years of training on the job. Duncan (1996) data from various resources on the effectiveness of two cognitive apprenticeship techniques, modeling and scaffolding, indicates that think aloud modeling was effective in increasing writing skills while scaffolding did not prove to be successful. Cash, Behrmann, Stadt, and Daniels (1997) conducted an experimental study where they found that cognitive apprenticeship was significantly more effective than traditional instruction in acquiring information, troubleshooting knowledge of air-conditioning, and diagnostic skills on automobiles within the settings of college automotive technology classes. And Pedersen and Liu (2002) claim that when students are not presented with good models of problem solving, they have to attain those skills through experimenting, which may result in learning undesirable knowledge and skills. They consider cognitive apprenticeship as a cure for this and use it in support of problem based learning. They studied modeling within problem based learning settings and found that modeling is effective to align students' work with expert ways of doing and enhance students' reasoning capabilities.

Exploration

With several positive examples and reports, and no significant detracting or unsuccessful reports found, there is reason to be optimistic about the efficacy of implementing cognitive apprenticeship in online training. Exploration in particular sounds premising in our case when the availability of online resources is mixed with the ability to manipulate the online learning environment to be created. Importance of pushing students to explore has been pointed out by several authors. Collins et al. (1989) describes exploration as "pushing students into a mode of problem solving on their own. Forcing them to do exploration is critical, if they are to learn how to frame questions or problems that are interesting and they can solve" (p. 481-482). Bonk and Wang (2001), for example, states that "encouraging learners to independently pursue investigations is the essential provision of an exploratory learning environment."

Secondly, to serve to IT Training and Education program's implied mission of making rapid products and making them as affective as possible, in this study, already existing workshop materials have been adopted to online settings. These existing materials consisted of instructional methods and strategies that can easily be associated with cognitive apprenticeship techniques. Removing them would reduce the usefulness of these materials. In such conditions, exploration manipulation was forming an excellent framework that was applicable.

Exploration in cognitive apprenticeship has not been studied, if not very little. Studies on exploration, independent of cognitive apprenticeship framework, refer to *discovery learning* in literature. In discovery learning, learners are presented with ready information. Discovery learning implements a learner centered approach to learning. The learning settings are usually represented by simulations. The learners therefore master this present information.

Literature on discovery learning reveals both effective and limitations of discovery learning. Stoddard (1985) studied styles of learning in an online help environment and denoted that tendency toward guided discovery learning was more preferred than structured learning. In support of this idea, various successes were handed over by several authors. For example Brylinsky (1995) studied undergraduate students in a required motor learning class and found that students in discovery laboratory settings scored higher than students in traditional learning settings in terms of laboratory report writing and, slightly, course examinations. Swaak (2001) found that effective discovery learning opportunities can be created with instruction that has certain qualities. There are studies that get our attention on downsides of discovery learning as well. For instance, in Hickey (1990) study, although the success of discovery learning is supported, it is added that guidance is needed to backing it efficiently. de Jong and van Joolingen (1998) collected quite a few of these findings that are successive and shortcoming aspects of discovery learning and categorized them into the following set of information. According to de Jong and van Joolingen (1998) discovery learning can successfully be attained if students are provided with assignments, direct access to information, model progression, and structured work environment. However, the discovery learning will not be as effective as traditional forms of learning environments as the discovery learning has the pessimistic potential in which students may have difficulty in the performance of hypothesis generation, experimenting the promoted behaviors, interpreting data and carrying out the learning without teachers' monitoring.

Context

Informed by the recognized success stories and research concerning cognitive apprenticeship, our focus was in applying this model to serve online skill-based computer workshops. The workshop model within IT Training and Education at Indiana University emphasizes a linear approach with guided activities typically culminating in a single completed product. This teaches specific skills with the intention of demonstrating best practices toward common larger projects. Since these face-to-face instructor-led workshops have a limited time-frame (three hours on average), there is little opportunity for deviation from the scripted material, branching to additional techniques related but secondary to the primary task, or student investigation of the tools or techniques.

By offering largely asynchronous online facilitated workshops, we recognized that the former time limitations would not necessarily apply. This would allow more opportunities for students to explore and delve deeper into the tools and concepts with the guided assistance of a knowledgeable instructor. Additional informational and educational resources could also be efficiently and economically made available through the Web. Within the traditional model of cognitive apprenticeship, exploration serves the purpose of allowing students a safe way to investigate and experiment. This concept has good face validity and apparent worth. It's easy to see why this would lead to expectations of deeper learning, greater long-term learning gains, and more satisfaction with the learning process. It can, however, require a good deal of additional effort on the part of instructional designers in crafting exploratory activities, on the part of instructors in providing direction and scaffolding for student exploration, and on the part of students engaging in these activities.

This research was intended to help shape the structure and focus of developing an online instructor-facilitated component of the IT Training and Education program. It was also hoped that the findings would have importance in adding to the literature of cognitive apprenticeship and research pointing to its effectiveness, particularly for online training but also potentially for all educational forums.

Research Questions

In the IT Training and Education model for development and delivery of online workshops, time and efficiency in the building and facilitation of these workshops will be a critical component. If methods and strategies such as exploration would not show significant benefits for students, then their inclusion in our model of rapid development would become counter-productive. Thus we wanted to put exploration to the test. Our hypothesis was that the inclusion of exploration activities in an online workshop for computer training would correlate with increased student learning, retention of learning, and student satisfaction.

Additionally, we wanted to gather data concerning time-on-task. Does exploration add significantly more time for students to complete a workshop?

Methods

Participants

Participants were recruited via listserv postings at Indiana University in early August of 2003. The lists covered multiple satellite campuses beyond the Bloomington campus. Participants were offered the chance to take the pilot online workshop "XHTML: The Basics" for free. Residential IT Training and Education workshops are offered free for students, faculty and staff pay $30 for each workshop, and participants outside the IU community may register for $60 (prices vary per workshop, but this is the standard pricing structure for the majority of workshops.)

Of the thirty-six who responded and enrolled, the vast majority were faculty and teaching assistants. The time of year (mid-August, 2003) may have played a large factor in these demographics as most students are away from school activities, staff and teaching faculty are between semesters, and many teaching faculty would have a need to put materials online for upcoming Fall courses.

These respondees were randomly assigned to one of two groups. The experimental group (Group A) had several exploratory activities built into the online workshop. Additionally, the instructor for this workshop was directed to encourage exploratory activity and, if appropriate, introduce new exploratory activities. The control group (Group B) had no additional exploratory activities in their online workshop and the instructor was instructed to not introduce new exploratory activities.

Procedures

The residential IT Training and Education workshop XHTML: The Basics was modified for Web delivery through two separate Web sites. One site (experimental) included additional exploration exercises and resources. In all, throughout the new sections of the workshop, there were six additional exploratory exercises built in to the material. There was one additional larger exploratory exercise alluded to in the material and directed by the instructor. This was a culmination exercise based on, at the students choice and according to their needs, the creation of a culmination Web site including all the elements covered in the workshop.

Two instructors were recruited from the IT Training and Education teaching staff. One of these instructors had previous experience with teaching an earlier online pilot workshop. This instructor was assigned to instruct the control group. A second instructor, new to the IT Training and Education staff, was assigned to the experimental group. Both instructors received an instructional manual and brief training session to cover the model of cognitive apprenticeship and example techniques.

Students were provided with a technical support contact to answer technical questions about the online environment itself beyond the scope of workshop materials. This technical support role handled connection and initial configuration issues and passed along technical problems related to the workshop Web sites themselves.

Communication and workshop interaction was accomplished using the Oncourse online course management system (http://oncourse.iu.edu), a system developed and maintained by Indiana University. Separate Oncourse environments were created for the two groups and students were added to the rosters for these according to their random assignment. Students were initially contacted by e-mail and told where and how to access their Oncourse environment. The first task they were assigned was the completion of a pretest. Students were given approximately two days to complete the pretest. At the end of that period, the pretest was disabled and students were given access to the course Web sites.

Students were given two weeks to complete the workshop from August 20 through September 2, 2003. At the end of the second week, a post-test and end-of-workshop survey were made available and students were encouraged to complete these.

Instruments

Both the pre-test and post-test consisted of 25 total questions: 10 multiple choice, 10 True/False, and 5 short coding questions. The multiple choice and true/false questions were worth one point each and the coding questions were worth varying points from 1 to 4 based on the number of specific elements students were being asked to demonstrate. These tests were written by the workshop author. The pre-test and post-test were online tests students completed in the Oncourse environment.

The end-of-workshop survey was based on an existing survey used by IT Training and Education for their residential workshops but modified and expanded for the sake of this online pilot study. This consisted of 22 questions in all, most with programmed scale responses but including several open-ended questions meant to gather greater detail in support of responses. The specific questions asked attempted to gain information from participants concerning demographics, feedback on instructors role, perceptions about motivation, satisfaction, learning, and time spent on the workshop. The survey was also an online survey completed in the Oncourse environment.

Our initial design also called for a follow-up survey and test of participants up to one month following completion of the workshops to evaluate perceptions of learning, satisfaction, and applicability as well as retention of learning. Time and resource limitations made this follow-up impossible to accomplish for this pilot test.

Results

Eighteen students were initially enrolled in each (control and experimental groups) workshop.

Pre-test/Post-test Completion and Scores:

In the experimental group, nine students completed the initial pre-test. The average score for this group was 34.3%. Eight students from this group also completed the post-test. Two who completed the pre-test did not complete the post-test, while one who did not complete the pre-test did complete the posttest. The average score for this group was 84.4%.

In the control group, thirteen students completed the initial pre-test. The average score for this group was 27.1%. Only four students from this group also completed the post-test, all of whom had also completed the pre-test. The average score for this group was 85.4%.

Survey Results

In both groups, the same individuals who completed the post-test also did complete the end-of-workshop survey (nine in the experimental group, four in the control group). Notable differences between the responses given between the two groups include the following.

When asked to agree that the presence of a facilitator was necessary for this workshop, on a scale of 1 (low) to 5 (high), the mean response from Group A was 3.55 (sd 0.95). The mean response from Group B was 1.75 (sd 0.43).

When asked if they felt the workshop was effective for learning (same 1 to 5 scale), Group A’s mean response was 4.44 (sd 0.49) while Group B’s mean response was 3.75 (sd 0.43).

When asked if they felt motivated to complete the workshop, the experimental group’s mean response was 4.11 (sd 0.56) while the control group’s mean response was 4.25 (sd 0.43).

When asked if they enjoyed taking the workshop, the experimental group’s mean response was 4.22 (sd 0.91) while the control group’s mean response was 4.25 (sd 0.43).

When asked if the topics covered in the workshop were helpful in learning the subject, the experimental group’s mean response was 4.66 (sd 0.47) while the control group’s mean response was 4.5 (sd 0.5).

When asked if the exercises in the workshop were helpful in learning the subject, the experimental group’s mean response was 4.55 (sd 0.49) while the control group’s mean response was 4.25 (sd 0.43).

When students were asked to estimate how much time, in two-hour increments, they spent working on this workshop, control group respondents were fairly evenly distributed under the 10 hour upper limit while experimental group participants responses were heavily concentrated in the eight to ten hour or more range. Responses are indicated in Table 1.

Table 1. Participant perspective of time spent completing the workshop.

Group A		**Group B**	
Response	*Time Increment*	*Response*	*Time Increment*
	Less than 2 hours		Less than 2 hours
	2 – 3.9 hours	1	2 – 3.9 hours
2	4 – 5.9 hours	1	4 – 5.9 hours
1	6 – 7.9 hours	1	6 – 7.9 hours
4	8 – 9.9 hours	1	8 – 9.9 hours
2	More than 10 hours		More than 10 hours

When asked if they felt the two week time frame for the workshop was sufficient, six of the respondents in the experimental group felt it was not enough time, while three from the experimental group felt it was just right. In the control group, two felt it was just enough time and two responded it was too much time.

Course Activity Statistics

Group A

Total Users: 22 (including 3 investigators and 1 tech support)
Chat Room Postings : 359
# of participants: 6 students, 1 instructor
Mail messages : total 47 messages sent by:
8 students (25 messages)
1 instructor (21 messages)
and 1 tech support (1 message)
Forum Postings:
Total Posts: 120, Total Read: 1023 (663, excluding those read by investigators)
Participants: 9 students (excluding instructor and investigators)
Lurkers: 3 (out of 9 students) – who read the posts but did not post any

Group B

Total Users: 24 (including 3 investigators and 1 tech support)
Chat Room Postings : 0
Mail messages : 0
Forum Postings:
Total Posts: 11, Total Read: 74 (41, excluding those read by investigators)
Participants: 8 students (excluding instructor and investigators)
Lurkers: 4 (8 students) – who read the posts but did not post any

Findings

While these results do not reflect adequate numbers to draw conclusions, they do point to some interesting indicators. The learning gain between groups was nearly identical. The control group, in fact, saw greater increase in test scores though not at a statistically significant level. Satisfaction with the workshops was nearly identical across the two groups. Long-term retention of learning could not be assessed since we were unable to conduct follow-up review.

Discussion

Clearly our hypothesis was not supported by these findings. While the results were problematic due to overall sample size exacerbated by drop-outs, we suspect that these findings would again be evidenced if this research were run again. Based on these results and our observations watching the study unfold, we would tend to believe that exploratory activities in and of themselves will not lead to greater immediate learning gains or levels of satisfaction, though we suspect that long-term learning would show improvements with more exploratory activities.

Limitations of the Study

Limitations of this study should be noted here. First, the findings of this study cannot be generalized to the population due to the small sample size. Although the descriptive analysis of the study revealed some notable differences between the two groups being examined, the investigators did not use or apply and referential statistics to the data because the sample size was not large enough for us to claim any statistical significance of those findings. Secondly, the investigators postulate that the instructors might have affected the results of the study. Although the investigators provided instructor training for both of the instructor, they exhibited different styles in their online facilitation skills. Specifically, the instructor for the experimental group used early and frequent probing questions and engaging dialogue to elicit and maintain student reaction. The instructor for the control group posted several initial messages but, after receiving no responses from students, did not actively attempt to elicit student interaction. Therefore, it is speculated that their different facilitative styles might have contributed to the differences in the level of learner participation in asynchronous online discussions between the two groups under study. Third, this study looked into only one component of the six techniques that an instructor can use to facilitate learning from a cognitive apprenticeship approach. Therefore, it is not known from this study how other cognitive apprenticeship strategies applied in the online courses under study might have contributed to the learning outcomes and learner satisfaction. Finally, the technology itself may have played a factor in student attrition. While the investigators had no direct evidence concerning why some participants did not continue, there were participant comments to Tech Support that indicated students were not familiar with the Oncourse utility and did not want to take the time to learn this new environment.

Additional Noteworthy Observations

Perhaps the most striking noted difference between these two groups is the amount of initial interest and subsequent drop-out on the part of participants. In the control group, thirteen of the initially enrolled eighteen students completed the pre-test, yet only four completed the post-test and end-of-workshop survey. In the experimental group, nearly all of the initial nine participants who completed the pre-test also completed the post-test and survey. This would cause us to speculate that, while active facilitation and exploratory activities do not promote better learning or satisfaction, they do play a role in engagement and motivation. This “humanization” may have implications for short online training workshops in attracting and maintaining participant interest and, subsequently, in attracting repeat customers.

The participant time spent on the workshop was also a surprise. Participants in the experimental group spent significantly more time engaged in the workshop than those in the control group. Yet the perceptions of the control group were split between the two-week time frame being adequate or being too long. In the experimental group, none of the participants felt the workshop time frame was too long, and the two-thirds majority felt the time frame was not long enough. Thus those in the exploration treatment did indeed spent much more time working on the workshop, but they were not dissatisfied with the time spent and our interpretation of their responses was that they were quite willing to continue the workshop. Implications for online training workshops may include the support for maintaining an extended online presence, building an online community to attract and keep the interest of participants.

Recommendations for Future Research

Another study of the online classes with a larger sample of students is warranted to yield results that are statistically significant, which will enhance the generalizability of the study. Results of such a study will be able to confirm or refute the results of the preliminary study presented in this paper. Also, since the results of this study suggest that the use of exploration activity contributed, in part at least, to active online interaction between and among participants of the online course, it could be worthwhile to investigate the effects of the use of the exploration activity on the promotion of higher-order thinking on the part of the student. Although there was no notable difference in the student's learning outcomes between the two groups in terms of short-term memory of learned skills in pre- and post- tests in this study, the analysis of the content of online discussion might be able to reveal the difference in the students' higher-order thinking levels between the two groups. In fact, the investigators believe that the application of a cognitive apprenticeship approach to instruction should contribute to deeper learning since the goal of the cognitive apprenticeship approach to teaching is to help students think like an expert, who is capable of applying his or her knowledge and skills to solve complex real-world problems. Therefore, a study of the effects of the cognitive apprenticeship approach to the student's deeper learning is recommended to investigate the practical implications of the cognitive apprenticeship method in teaching and learning.

References

Bonk, C. J. (1998). Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools. In C. J. Bonk and K. S. King (Eds.), *Electronic collaborators learner-centered technologies for literacty, apprenticeship, and discourse* (pp. 25-50). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Bonk, C. J., and Wang, F.-K. (2001). A design framework for electronic cognitive apprenticeship. *JALN, 5*(2).

Brown, J. S., Collins, A., and Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Brylinsky, J. (1995). Authentic Discovery Laboratories in Motor Learning.

Casey, C. (1996). Incorporating cognitive apprenticeship in multi-media. *Educational Technology Research and Development, 44*(1), 71-84.

Cash, J. R., Behrmann, M. B., Stadt, R. W., and Daniels, H. M. (1997). Effectiveness of cognitive apprenticeship instructional methods in college automotive technology classrooms. *Journal of Industrial Teacher Education, 34*(2), 29-49.

Cheng, C.-C. (1995). Proactive guidance in computer-assisted language learning.

Collins, A. (1991). Cognitive apprenticeship and instructional technology. In L. Idol and B. F. Jones (Eds.), *Educational values and cognitive instruction: Implications for reform* (pp. 121-138). Hillsdale, NJ: Lawrence Erlbaum.

Collins, A., Brown, J. S., and Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator, 15*(3), 6-11, 38-46.

Collins, A., Brown, J. S., and Newman, S. E. (1989). Cognitive apprenticeship: teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

de Haan, M. (2002). Distributed cognition and the shared knowledge model of the mazahua: A cultural approach. *Journal of Interactive Learning Research, 13*(1-2), 31-50.

de Jong, T. v. J., Wouter R. (1998). Scientific discovery learning with computer simulations of conceptual Domains. *Review of Educational Research, 68*(2), 179-201.

Derry, S. P., and Lesgold, A. M. (1996). Toward a situated social practice model for instructional design. In R. Calfee and D. C. Berliner (Eds.), *Handbook of Educational Psychology* (pp. 787-806). New York: Macmillan.

Duncan, S. L. S. (1996). Cognitive apprenticeship in classroom instruction: Implications for Industrial and Technical Teacher Education. *Journal of Industrial Teacher Education, 33*(3), 66-86.

Elliott, C. B. (1994). *Pedagogical reasoning: Understanding teacher decision making in a cognitive apprenticeship setting.* Unpublished doctoral dissertation, Texas Woman's University.

Hendricks, C. C. (2001). Teaching causal reasoning through cognitive apprenticeship: What are results from situated learning? *Journal of Educational Research, 94*(5), 302-311.

Hickey, D. T. P., Anthony. (1990). Effects of generative video on students' scientific problem posing.

draft.
Horiguchi, T. H., Tsukasa. (2000). A framework for creating counterexamples in discovery learning environment.
Johnson, S. D., and Fischbach, R. M. (1993). Teaching problem solving and technical mathematics through cognitive apprenticeship at the community college level.
Kashihara, A. K. O., Reinhard Rashev, Rossen Simm, Helmut. (2000). A cognitive load reduction approach to exploratory learning and its application to an interactive simulation-based learning system. *Journal of Educational Multimedia and Hypermedia, 9*(3), 253-276.
Keller, B. R., Charles M. (1985). A comparison of three instructional presentation formats. IDD&E working paper No. 6.
Klein, A. (1991). All about ants: Discovery learning in the primary grades. *Young Children, 46*(5), 23-27.
Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge, UK: Cambridge University Press.
Leont'ev, A. (1981). The problem of activity in psychology. In J. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 37-71). Armonk, NY: M. E. Sharpe.
Luira, A. (1979). *The making of mind: A personal account of Soviet psychology*. Cambridge, MA: Harwerd University Press.
Mitchell, R., and Myles, F. (1998). Sociocultural perspectives on second language learning. In R. Mitchell and F. Myles (Eds.), *Second language learning theories* (pp. 144-162). London: Arnold.
Nicaise, M. (1997). Computer-supported apprenticeships in math and science. *Journal of Computers in Mathematics and Science Teaching, 16*(4), 443-465.
Pedersen, S., and Liu, M. (2002). The effects of modeling expert cognitive strategies during problem-based learning. *Journal of Educational Computing Research, 26*(4), 353-380.
Practitioner Research Briefs, 1998-1999 Report Series. (1997).
Rose, D. (1995). *Apprenticeship and exploration: A new approach to literacty instruction (Literacy Research Paper 6)*. New York: Scholastic.
Stein, D. (1982). Situated learning in adult education. ERIC Digest No. 195.
Stoddard, M. L. (1985). Learning styles and embedded training: A case study.
Sugrue, B. M. T., Rex A. (1987). Effects of discovery and expository instruction on recall and transfer of procedural knowledge: Interactions with learner aptitude.
Suomala, J. A., Ari Alajaaski, Jarkko. (2000). Students' problem-solving in a complex technology-based learning environment.
Swaak, J. d. J., T. (2001). Discovery simulations and the assessment of intuitive knowledge. *Journal of Computer Assisted Learning, 17*(3), 284-294.
Teles, L. (1993). Cognitive apprenticeship on global networks. In L. M. Harasim (Ed.), *Global Networks: Computers and International Communication* (pp. 271-281). Cambridge, MA: MIT Press.
Vygotsky, L. S. (1979). Consciousness as a problem in the psychology of behavior. *Soviet Psychology, 17*, 3-35.
Wilcox, R. T. (1987). Rediscovering discovery learning. *Clearing House, 61*(2), 53-56.
Wilson, B. G., Jonassen, D. H., and Cole, P. (1993). Cognitive approaches to instructional design. In G. M. Piskurich (Ed.), *The ASTD handbook of instructional technology* (pp. 21.21-21.22). New York: McGraw-Hill.

Technology Loving Psychologists: How are they *really* using the technology?

Xornam Apedoe
University of Georgia

Connie Varnhagen
University of Alberta

Abstract

The influx of technology into higher education over the past few years has been phenomenal. Yet the question remains, how is the technology being used for instruction, and more specifically in the instruction of psychology? A survey was developed, designed to focus on: (a) levels of instructional technology use, (b) instructor characteristics related to use, and (c) institutional characteristics that promote or inhibit use. The survey was distributed by mail to instructors of introductory psychology at Canadian universities in 200 and has been updated and distributed by Email to instructors of introductory psychology at U.S. and Canadian universities in 2003. Results from the 2000 survey show that in the instruction of introductory psychology, perceived usefulness of instructional technology is much higher than actual use. Actual use of instructional technologies is highly influenced by instructors' experience with technology. Lack of support is most frequently stated as being a barrier to using technologies, while having a support center and release time for development are stated as the two factors that may increase use of instructional technologies. The 2003 results will be presented and compared with the 2000 results in the presentation.

Technology Loving Psychologists: How are they really using the technology?

Delivering and accessing high quality undergraduate education has become much more difficult today than at any other time in history. Higher education is currently being forced to deal with difficult issues such as rising tuition costs, a decreased focus on teaching and student advising by faculty, outdated curriculums that are insensitive to the needs of individual learners and the realities of the workplace, increased numbers of non-traditional students, ageing facilities, and decreased government funding (Van Dusen, 1998; Surry & Land, 2000). Many university administrators envision technology as providing the solutions to many of the problems of higher education, being both cost-effective and innovative (Surry & Land, 2000). Technological innovations are believed to make "the academy more accessible, more affordable, and more effective" (p. 59, Van Dusen, 1998). More importantly, technology is now viewed as being a potentially valuable tool for educational reform (Surry & Land, 2000) enhancing both teaching and learning. Thus we are now seeing increasing amounts of access to innovative technologies for both students and faculty in higher education (Surry & Land, 2000).

Deleted: :

Although the field of educational psychology has long been concerned with issues related to research on teaching and learning, many psychology instructors have always demonstrated interest in addressing such issues. National and international organizations such as The Society for the Teaching Psychology (http://teachpsych.lemoyne.edu/), and the Learning and Teaching Support Network for Psychology (http://www.psychology.ltsn.ac.uk/) are actively involved in promoting exceptional teaching and learning of psychology at the undergraduate level. Publications such as *Teaching of Psychology*, promote the sharing of research on innovative instructional practices as well as the sharing among colleagues of "best" teaching practices. With the increasing prevalence of technology in institutions of higher education, more research is needed to examine issues related to the use of technology in the instruction of undergraduate psychology.

Numerous studies have examined faculty members' use of technology for instruction in higher education, but very few have focused on the use of technology in psychology instruction. There are numerous Web-based and computer-supported resources that deal specifically with psychology content, however there is very little research examining if and how psychology instructors are using these resources. Thus, there are three issues of primary concern here. First, a level of use needs to be established. How much, and what kinds of technology are used for instruction? What other instructional methods are commonly employed? Second, of interest are factors that promote or inhibit the use of instructional technology. What institutional characteristics commonly promote or inhibit the use of instructional technology for psychology instructors? Third, it is important to highlight some of the characteristics of individuals that use or do not use instructional technology. Are there any individual characteristics that can identify frequent users of instructional technology?

Previous Surveys

Technology use by higher education faculty from a number of disciplines was examined in studies conducted by Spotts and Bowman (1995), and research groups from Academic Technologies for Learning (ATL) and Learning Support Systems (LSS) (1996). Spotts and Bowman (1995) examined the levels of knowledge and experience with instructional technologies of faculty at Western Michigan University. The survey examined: (a) instructors' knowledge and experience with newer technologies, (b) instructor characteristics related to instructors' knowledge and experience, and (c) instructors' perceptions and attitudes towards technology. The survey conducted by ATL and LSS (1996) provides a Canadian perspective on the use of instructional technologies in a higher education institution. Their survey assessed the (a) attitudes, (b) use, and (c) concerns, of academic staff at the University of Alberta, Canada, regarding instructional technologies.

Respondents in Spotts and Bowman's (1995) study reported being most knowledgeable and experienced with the commonly used technologies of audio, video, film, and word processing. Fewer faculty members reported high knowledge for computer technologies such as spreadsheets, Email, and computer assisted instruction. A high number of faculty members reported having no experience with computer conferencing, presentation software, and multimedia. Results also showed that there were gender differences in knowledge and experience. Males reported higher levels of knowledge and experience with newer technologies (multimedia, distance learning, CAI, Email, computer conferencing and presentation software) than their female colleagues reported. Spotts and Bowman cautioned that before their findings can be generalized across the higher education community "future research on other populations and types of institutions is needed…[and] future studies about the influence of discipline…are needed"(p. 63).

A survey conducted in 1996 by ATL and LSS at the University of Alberta, a large Canadian university with a mandate to use instructional technology in teaching and learning, assessed the attitudes, use, and concerns of academic staff at the University of Alberta regarding instructional technologies. Some key findings regarding use and experience included: (a) faculty rating their use of technology as being moderately effective, (b) ratings of computing skill was higher for technologies that faculty typically used such as word processing, Email, and database searches and lower for instructional technologies such as presentation software, Web page creation, and course authoring software, (c) more experienced technology users tended to rate the importance of technology higher, and (d) younger faculty tended to rate the importance of technology higher. The highest rated barriers to technology use were (a) institutional or departmental funding, and (b) time required for learning new technologies. Although many respondents indicated wanting more evidence of the advantages of using the newer instructional technologies, many still believed that new technologies will enhance classroom instruction. The ATL and Learning Support Services survey also found notable differences in responses from the different university faculties. In support of Spotts and Bowman's (1995) findings, this suggests that a closer look at the use of, experience with, and attitudes towards instructional technology within different faculties or disciplines is required.

A survey to examine the use of computers in psychology instruction was conducted by Hornby and Anderson (1994). Hornby and Anderson's survey focused on the use of technology by psychology instructors who were skilled in computer use. The goal of their survey was to examine the individual and institutional characteristics that relate to the use of computers in instruction. To this end, they examined the individual and institutional characteristics that may be useful in differentiating between levels of technology users, as well as those characteristics that may be associated with promoting or inhibiting use of technology for instruction. Hornby and Anderson also explored factors that individuals would consider important for increasing use of technology for instruction.

Hornby and Anderson's (1994) survey results showed that although a majority of faculty members consider themselves dedicated users of computers for instructional purposes, two-thirds reported using the computer for instruction for less than 3 hours per week. Regarding the individual factors that relate to instructional use of computers, Hornby and Anderson (1994) found a positive relationship between academic rank and use of computers for testing. Years of teaching positively correlated with instructional computing use, while instructors holding a masters degree reported using computers in instruction more than those who held a PhD. Non-tenured faculty members were less likely to report using computers for testing purposes. There were few institutional characteristics reported to influence the levels of instructional technology use. Use of instructional technology was more prevalent at undergraduate institutions that had a stronger emphasis on teaching than on research. Although previous studies (Castellan, 1982) found a relationship between the size of the student body and the use of computers in instruction, none was found in Hornby and Anderson's survey (1994). Finally, over 50% of respondents indicated that release time for development and better software would both be important factors in encouraging their use of technology in their teaching. Evidence of the effectiveness of instructional technology, and positive evaluation for tenure or promotion, were relatively less important factors in encouraging use of instructional technology for respondents.

Present Study

To expand on the current literature on instructional technology use in higher education, and as a follow up to the study conducted by Hornby and Anderson (1994), we developed a survey to assess common instructional practices used in psychology. To identify factors that can be influential in improving the quality of instruction of introductory psychology, the survey focused specifically on those instructional practices commonly used in the teaching of introductory psychology. The survey measured the amount of instructional technology used in introductory psychology, as well as examined reasons why instructional technologies are, or are not used in instruction. All instructors of introductory psychology at major English speaking Canadian universities were invited to complete the survey, regardless if they did or did not use instructional technology. To prevent any sample biases, administration of the survey was via paper and pencil instead of using Email or the Web.

The goals of the survey were to: (a) establish a level of instructional technology use in introductory psychology, (b) examine the individual characteristics of instructors that relate to instructional technology use, and (c) examine the institutional characteristics that promote or inhibit instructional technology use. The individual characteristics examined included age, gender, academic rank, total number of years teaching, total number of years teaching introductory psychology, and experience with technology. The institutional characteristics examined included geographical location, university population size, and departmental emphasis on research and teaching.

Method

Participants

Respondents to the 2000 survey included sessional instructors, graduate students, post-doctoral fellows, and professors that were currently (1999/2000 academic year) teaching introductory psychology at any English speaking Canadian University. To determine the number of instructors who were currently teaching introductory psychology, information was obtained directly from departments and department Web sites. A corresponding number of surveys were then distributed to each institution. Responses were voluntary, and participants were assured that their responses would remain anonymous (i.e. no efforts would be made to identify them by institution).

We are completing data collection for the 2003 survey and these results will be presented at the conference.

Materials

A 45-question survey was developed based on previous questionnaires and surveys. The Department of Psychology at the University of Alberta reviewed the survey for ethical acceptability. The survey consisted of seven main sections covering a broad range of topics including: General Attitudes, Course Demographics, Instruction, Technology Skill, Use/ Non-Use of Technology in Instruction, University and Department Resources and Demographics, and Personal Demographics.

Procedure

A package of surveys was distributed to the Chair or Department Head of all the English speaking Canadian universities listed on the CPA Website (http://www.cpa.ca). Included in the survey package was a cover letter addressed to the Chair, or Department Head, requesting surveys be distributed to any professor, graduate student, post-doctoral fellow, or sessional instructor who was currently (1999/2000 academic year) teaching introductory psychology. Two mailings were conducted, with 176 surveys distributed in the first mailing. Two institutions responded with requests for more surveys, leading to the distribution of 8 additional surveys in a second mailing. In total, 184 surveys were distributed to 43 different universities. Attached to each survey was an addressed stamped envelope in which the completed survey could be returned. A follow-up Email sent to Chairs and Department Heads requested they remind those instructors who received the survey to complete and return them.

Results

Sixty-one surveys (33%) were completed and returned. Response rates for each geographical region of Canada (BC, Prairies, Ontario, Quebec) was above 35%, based on the number of surveys distributed to institutions within each region. The only exception was in the Maritimes, which had a response rate of only 21%. Of the total number of responses received, approximately 16% came from institutions in British Columbia, 32% from the Prairie provinces, 31% from Ontario, 3% from Quebec, and 18% from the Maritime provinces.

Respondent Demographics

Sixty-seven percent of respondents were male and 30% of respondents were female. Three percent of respondents did not indicate their gender. Age of respondents ranged from 29 to 68 years, with a mean age of 46 years (SD = 10.4). The distribution of respondents' academic rank was fairly equal with 25% non-tenure faculty, 23% assistant professors, 27% associate professors, and 25% full professors. Teaching experience ranged from novice (1 year) to very experienced (39 years). The mean number of years of experience for teaching introductory psychology was 12 years (SD = 10.3).

Respondents had a wide range of experience with the different technologies. Respondents were most experienced with Email, with 92% of respondents rating themselves as being experienced or very experienced. Following Email, respondents were most experienced with presentation software (i.e. PowerPoint) with 54% rating themselves as experienced or very experienced. Surprisingly, only 44% of respondents rated themselves as being experienced or very experienced with the Web, with 36% rating themselves as being not very experienced or only slightly experienced. The majority of respondents rated themselves as not very or slightly experienced with multimedia or interactive programs (79%), programming (72%), and Web design/maintenance (64%).

Deleted:

Establishing Level of Technology Use

Respondents were asked to indicate how frequently they used various technologies in their instruction on a Likert 5-point scale as never (1), once per semester (2), once per month (3), weekly (4), or every class (5). The vast majority of instructors use conventional technologies (i.e. overheads, chalkboards, and slides) weekly or every class (82%). Audio and/or video (56%) as well as presentation software such as PowerPoint (46%), are also used weekly or every class by a large number of instructors. Only a small number of instructors use other technologies, such as course Web pages (31%), Web links (22%), multimedia software (17%), asynchronous discussions (10%) and synchronous discussions (2%) weekly or every class.

Respondents also rated how useful they believed specific technologies are for instruction on a 5-point Likert scale that ranged from (1) not very useful to (5) very useful. Not surprisingly respondents felt that conventional technologies were the most useful, with the vast majority rating them as being useful or very useful (81%). The vast majority of respondents also rated audio and/or video (80%), and presentation software (72%) as being useful or very useful. Except for the conventional technologies, the perceived usefulness of all the other technologies was rated much higher than their reported levels of use. Table 1 shows the perceived usefulness of each technology compared with reported levels of use for the corresponding technology.

Table 1 *Use and Perceived Usefulness of Instructional Technologies*

Technology	Perceived Usefulness(%) (Useful to Very Useful)	Use (%) (Weekly to Every Class)
Conventional (slides, chalkboards, overheads)	81	82
Audio/ Video	80	56
Presentation Software	72	46
E-mail	65	63
Course Web Page	48	31
Web Links	38	22
Multimedia/ Interactive Software	37	17
Asynchronous Discuss	18	10
Synchronous Discuss	8	2

Instructor Characteristics Related to Technology Use

Not surprisingly, experience with a specific technology influences the use of the specific or related technology. Experience with technologies such as presentation software, the Web, Web design/ maintenance, and multimedia were all positively correlated with use of the same or related instructional technologies. Table 2 presents the association values between experience with and use of each technology.

Table 2 *Association between Experience and Use of Instructional Technologies*

Experience	Use	Kendall's Coefficient (W)
Presentation Software	Presentation Software	0.56*
Web	Web Links	0.52**
Course Web Page	0.41**	
Web Design	Course Web Page	0.55**
Multimedia	Multimedia	0.62**

* Significant at the $p < 0.05$ level
** Significant at the $p < 0.01$ level

Age is also an instructor characteristic that influences the use of technology for instruction. Although synchronous and asynchronous discussions were used by a small number of instructors, an interesting negative association was found between the use of these technologies and instructors' age (synchronous discussion: $W = -0.29$, $p < 0.05$; asynchronous discussion: $W = -0.25$, $p < 0.01$).

Institutional Characteristics Related to Promotion and Inhibition of Technology Use

When asked what factors might increase their use of technology, the majority (56%) of respondents reported that release time for development would be an important factor. A majority (51%) of instructors also reported that having a support center might increase their use of technology. Evidence of the effectiveness of instructional technology (38%), better software (30%), and counting towards tenure and promotion (30%) were considered by fewer instructors as being factors that would increase their use of technology in teaching introductory psychology.

Only one institutional characteristic related to respondents' indication of a factor that may promote their use of instructional technologies: geographic location. For instructors from institutions located in British Columbia, better software was a factor that may promote their use of instructional technologies ($V = 0.30$, $p < 0.05$).

Respondents were also asked to indicate factors that act as a barrier to their use of technology in instruction. Lack of support (62%) as well as lack of funding for technological innovations (61%), were indicated by a majority of instructors as being barriers to instructional technology use. Lack of knowledge was only indicated by 39% of instructors as being a barrier to their use instructional technology. Lack of incentives (36%) and lack of evidence of the effectiveness of instructional technology (36%), were also not considered barriers by a large number of instructors. Three percent of respondents indicated they had no barriers to using of instructional technology.

Geographic location of the institution was associated with indicating "lack of knowledge" as a barrier to use of instructional technologies. Instructors from institutions located in BC ($V = 0.26$, $p < 0.05$), the Prairies ($V = 0.27$, $p < 0.05$), and Ontario ($V = 0.39$, $p < 0.05$) indicated "lack of knowledge" as being a barrier to their instructional technology use. University population size was related to indicating "lack of funds" as being a barrier to instructional technology use ($V = 0.25$, $p < 0.05$).

Discussion

Overwhelmingly, instructors of introductory psychology continue to rely on the familiar technologies of overheads, chalkboards and slides. Integration of newer technologies such as course Web pages, Web links and multimedia software appears to be minimal. Instructors' experience with technology appears to act as a mediating factor in the amount and types of technology used. The limited integration of new technologies into instruction in introductory psychology may be explained by the reported amount of inexperience instructors have with these new technologies. These results support Spotts and Bowman's (1995), and ATL and Learning Support Systems (1996) findings that the technologies used by a majority of instructors are the technologies with which they are most experienced. With less than 50% of respondents in this study rating themselves as experienced or very experienced with the Web, it is no surprise that the use of Web related instructional technologies is relatively low.

While experience with technology emerged as the predominant individual instructor characteristic that influences respondents' use of technology for teaching psychology, geographical location and university population size appeared to be the only institutional characteristics related to the factors cited as influencing the promotion or inhibition of technology use. For respondents from all regions of Canada except for Quebec and the Maritime provinces, lack of knowledge was indicated as being a significant barrier to their instructional technology use. Therefore, increasing the knowledge and experience level of instructors at institutions in these regions may increase their instructional technology use. For universities with a larger student population, lack of funds is a concern of instructors that needs to be addressed.

Implications and Recommendations

Despite the apparent lack of integration of new technologies in the teaching of introductory psychology, anecdotal evidence suggests that there are a number of psychology instructors who are developing educational products and integrating technology into their instruction in innovative ways. In response to increasing class sizes and a desire to maintain a high quality standard for teaching, some psychology instructors are utilizing Web and computer-based educational software to extend students' learning experiences outside of the classroom, while other instructors make use of new technologies for enhancing the in-class learning experience. Educational software such as the Brain Study Module, developed in the Instructional Technology Lab in the Psychology Department at the University of Alberta, has been developed to expose students to concepts related to functions of the brain, and is currently used within a number of psychology courses across institutions (the module can be found on the WebCT Psychology Community page at http://www.webct.com/psychology/home/). Psychology instructors are also using technology within the classroom to enhance their teaching of psychology content. For example, Dr. Michael Atkinson, at the University of Western Ontario, is able to effectively teach an introductory psychology class of 1200 students with the help of technology. By using audio, video, and computer software, he is able to create an interactive learning environment for his students. With such exemplary examples of technological innovations that psychology instructors are making use of, it is important that research findings are used to inform future policies and decisions regarding the integration of technology into undergraduate psychology instruction.

Based on the results from this survey and the relevant literature, some common themes arise that can be used to structure recommendations to promote or increase the use of instructional technologies in the instruction of introductory psychology at Canadian universities. If institutions want to encourage the use of instructional technologies, they need to increase instructors' experience and technological skill level. Institutions may want to start by providing support centers and staff knowledgeable in the implementation of technologies in instruction. This can provide instructors with support in the appropriate use and integration of technologies, knowledge of the technologies available, and most importantly increase their skill and experience level with technologies. Institutions located in British Columbia, the Prairies, and Ontario should pay special attention to this issue because for respondents from these locations, lack of knowledge is a significant barrier to their use of instructional technologies.

Instructors also indicated that release time for development may increase their use of technology, which was also indicated by a majority of respondents in Hornby and Anderson's (1994) study. Therefore, institutions may also want to begin to develop policies that provide instructors with release time to develop appropriate technological innovations. The time commitment involved in developing pedagogically sound instructional materials using technology is immense. With the pressures and constraints on time most academics face today, learning how to use a technology and then developing instructional materials may not be possible without release time from some of their normal duties.

Conclusion

The present study should be viewed as a starting point for future research related to the use of technology in the teaching and learning of introductory psychology. Future research should continue to examine psychology instructors changing and evolving attitudes and experience with technology and the subsequent influence this has on the teaching of introductory psychology. Further research should also examine how psychology instructors' use of technology in teaching supports their beliefs about teaching and learning.

Acknowledgements

This research was conducted in the Department of Psychology at the University of Alberta. Funding was provided by LEE III and the Faculty of Science at the University of Alberta.

References

Academic Technologies for Learning & Learning Support Systems. (1996). *Enhancing the university's learning system through technology: A survey of academic staff.* Edmonton, AB: University of Alberta, ATL & Learning Support Systems.

Castellan, N. John, (1982). Computers in psychology: A survey of instructional applications. *Behavior Research Methods & Instrumentation, 14* (2), 198-202.

Hornby, P. A., & Anderson, M. D. (1994). Computer use in psychology instruction: A survey of individual and institutional characteristics. *Behavior Research Methods, Instruments, & Computers, 26,* 250-254.

Spotts, T. H., & Bowman, M. A. (1995). Faculty Use of Instructional Technologies in Higher Education. *Educational Technology, 35,* 56-64.

Van Dusen, G. C. (1998). Technology: Higher Education's Magic Bullet. *Thought & Action, 14* (1), 59-67.

Virtual Solar System: How do we know if students are learning?

Xornam Apedoe
Beaumie Kim
Kenneth E. Hay
University of Georgia

Introduction

Introductory astronomy is one of the main undergraduate courses offered in higher education institutions that many students enroll in to meet science credit requirements. It is certainly irrelevant to most academic majors, and memorized facts for quizes are quickly forgotten, as it may be their last science course (Skala, Slater & Adams, 2000). Introductory astronomy courses typically include a laboratory component as a way to provide students with a more hands-on experience of astronomy. Different approaches for science learning have been suggested and implemented for laboratory experiences as “real” science laboratories change their practices, as is the case in astronomy, where many laboratories have completely switched to doing computer experimentations (Marschall, 2000). Many astronomy educators suggest that in order to develop a deeper understanding of the subject, targeted concepts should be addressed for extended time using appropriate strategies (Prather, Slater, & Offerdahl, 2002).

Over the past 6 years, we have continued a reform effort for an undergraduate astronomy laboratory course, which utilizes unique strategies for teaching targeted concepts. The Virtual Solar System (VSS) project, funded by the National Science Foundation, combines modeling-based inquiry and a set of 3D model construction tools to enhance students’ understanding of astronomy. In the past decade numerous studies have shown that students have difficulties understanding basic astronomical phenomena (Wandersee, Mintzes & Novak, 1990). With the Virtual Solar System (VSS) software, students are able to build 3D models of the solar system to help them study and understand basic and complex astronomical concepts.

VSS Software

The Virtual Solar System software is a virtual reality modeling tool with which learners build their own solar system models and simulate them. It is a computer-based three-dimensional environment where learners can navigate through the created space. Figure 1 shows an overhead view of the Earth (at the center) and the Moon (selected box) orbiting around the Earth: the big circle represents the Moon’s orbital plane; the bottom right window shows the properties of the selected object (i.e., Earth’s Moon); the top left window shows the view from a different perspective (i.e., looking at the Earth in front); the middle left window controls the model to run, accelerate, decelerate, pause, stop, and jump forward and backward (e.g., jumping a year ahead or going back a year).

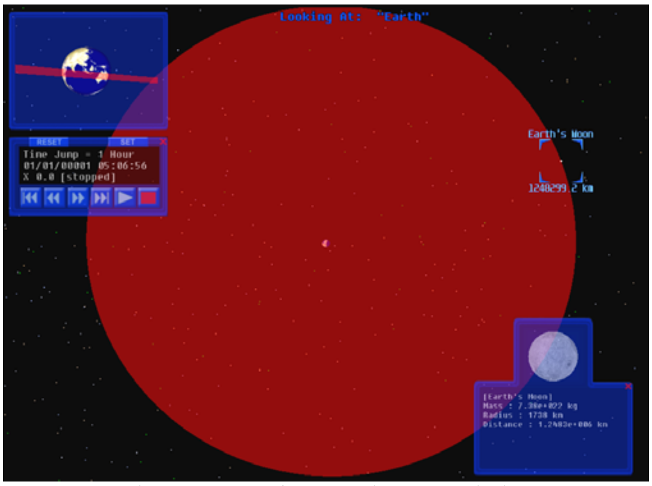


Figure 1. VSS software Basic Screen Display

Two Classes with Modeling-Based Inquiry

Modeling-based inquiry is a specific pedagogy focused on simulating the practice of scientists’ use of

computational models. Students use their models as a source of data collection to answer inquiry questions in the laboratory course. Students start with an inquiry question, plan and create a model, validate the model, address the inquiry question, explore and visualize data from the model, develop conclusions, and report findings.
We have developed and tested different conditions of implementing this pedagogy in a laboratory course setting addressing the same concepts of eclipses, seasons, phases, orbits, and time: constrained conceptual coherence condition and historical recapitulation condition.

Constrained Conceptual Coherence (CCC) Condition. As typified in conventional curricular models this condition provides questions that are logically organized throughout the course. The exercises start with easier questions and tasks and progresses to difficult ones. Students are directed to use certain tools provided within the VSS software. The seven units of the course include orbits, time, phases I, phases II, eclipses, seasons, and student choice.

Historical Recapitulation (HR) Condition. The students in the class that utilizes this condition are provided with the questions and tools that follow the historic development of astronomy knowledge. Wandersee (1986) suggests that learners benefit from addressing scientific domains by recapitulating the path through which scientific concepts are initially formulated. The historic path is mainly moving from the geocentric view (i.e., seeing the Earth as the center of the universe) to heliocentric view (i.e., the Sun as the center of our solar system) by finding the inconsistency of the observation in the sky. Specific features and tools, such as circular orbits, sky constellations, compass, landing on the surface of a planet, are provided for this condition so that students can simulate the experience of the scientist looking at the sky from the surface of the earth. The six units of this condition include time and motion, seasons, eclipses, planet motions, phases, and orbits.

Research Goals and Questions

The overarching focus of the ASTR 1010L course is to help students simulate the practices of scientists, particularly astronomers, and how they use computational models. Thus, the overall research goals for the VSS project are to examine the emergence, diffusion and evolution of students' astronomy knowledge and ability to simulate scientific practices. To gain a complete picture of the emergence, evolution and diffusion of students' knowledge, we need to combine rigorous qualitative observations with objective measures of conceptual knowledge. Qualitative observations focus on the process of inquiry learning, while objective measures assess students' understanding of basic astronomy concepts. The focus of this study is to assess students' knowledge of concepts that are the basis of the field of astronomy. In particular we are interested in the influence of pedagogy on objective measures of astronomy concept knowledge. The purpose of this research is to examine changes in students' astronomy knowledge as a function of format of the course. Our primary research questions are:

- Does use of modeling-based inquiry enhance students' understanding of astronomy concepts?
- Do different conditions of modeling-based inquiry (HR and CCC) have different influences on students' knowledge of astronomy?

Methods

Participants

This research takes place in a one-credit introductory astronomy laboratory course (ASTR 1010L). Students meet for 2 hours per week, for 15 weeks over a semester. Using a modeling-based inquiry pedagogical approach, students typically work collaboratively in triads to build models of the solar system with the VSS software. The instructor assesses student learning mainly through students' written reports and oral presentations. In the spring of 2003, there were two sections, based on the two conditions (CCC and HR as described above) of the ASTR 10100L.

Materials

The items used on the pre- and post-test are a compilation of items from two previously developed and validated tests: Project STAR-Astronomy Concept Inventory (Sadler, 1992) and the Astronomy Diagnostics Test (Zeilik, 1998). Items were developed to address concepts not covered in the Project Star and ADT inventories that are addressed in the ASTR 1010L. The tests were developed to serve two purposes. Firstly, the pretest serves to provide a baseline measure of students' initial understanding and misconceptions related to astronomy knowledge, which can be used as a starting point for tracing the emergence, evolution, and diffusion of students' astronomy understanding. Secondly, the pre- and post-test serves to assess the presence of common misconceptions. Thus, the test items were constructed using common misconceptions as answer alternatives. The pre-and post-test contain 53 questions that address the following astronomy concepts: eclipses, seasons, phases, orbits, and time. The pre-test has an additional 13 questions that solicit demographic information from respondents.

Procedure

Assessment methods are continually being developed and refined throughout this project. Currently a pre-post test design, using objective multiple-choice test items is being used. The pre-test was administered by a researcher to all participating students during the first class meeting to establish a baseline of students' astronomy knowledge prior to any instruction. The test was not shared with the instructor at any point to prevent any chances of 'teaching to the test.' The post-test was administered by the same researcher during the last class of the semester. Students were given as much time as needed to complete the test, although students typically finished within 45 minutes.

Results

Respondent Demographics

Pre and post-test data was collected from a total of twenty-four students enrolled in the two sections of ASTR 1010L in the spring of 2003 (11 from HR condition and 13 from CCC condition). The majority of students were under 23 years of age (95.8%), and the distribution of genders was fairly equal in each class with females comprising 54.2% of the students and males 41.7% (there was a non-response rate of 4.1%). Only a small number of students (8.3%) were Science, Engineering or Architecture majors. A large number of students indicated that their major or current area of focus was in Humanities, Social Science or the Arts (41.7%), Business (20.8%) or other areas of interest (20.8%). Finally, the majority of students were currently enrolled in the lecture section of the Astronomy class (58.3%), although a significant number of students had previously completed the course (37.5%). This survey question had a non-response rate of 4.2%, but the lecture section is a prerequisite course that students either take it with the lab or prior to taking the lab.

Pretest Results

The pretest provided a baseline assessment of students' initial knowledge of concepts related to eclipses, seasons, phases, orbits, and timekeeping. Students in both the Historical Recapitulation (HR) and the Constrained Conceptual Coherence (CCC) classes had limited astronomy knowledge related to the aforementioned concepts at the beginning of the semester. Students in the CCC class had a mean score of 26. 5, or 49.9% (SD = 7.92) while students in the HR class achieved a mean score of 24.5, or 46.2% (SD = 7.87). An analysis of variance of performance on the pretest indicated there were no significant differences in students' overall knowledge of astronomy concepts [F (1, 25) = 0.41, p = 0.53). Students in both classes also demonstrated similar levels of knowledge on questions specifically related to astronomy concepts such as eclipses, phases, orbits, and timekeeping (see Table 1).

Comment: check this – how do you really write this up?

Table 1. *Pretest Descriptives by Concept*

	Mean Score	
	HC	CCC
Eclipses	5.8	4.1
Seasons	4.9	7.0
Phases	2.6	2.9
Orbits	6.1	6.8
Timekeeping	2.2	1.8

Posttest Results

Students' in both the HR class and the CC class showed improved performance and knowledge of astronomy concepts on the posttest as compared to the pretest. The mean score on the posttest in the HR class was 30.6, or 57.6% (SD = 7.46), while the mean score achieved by students in the CC class was 36.4, or 68.7% (SD = 5.82). However, this improvement was only statistically significant for students who were participants in the CC class [F (1, 24) = 14.6, p = 0.00).

Comparison of Posttest Results for HR and CCC classes

Although there were no significant differences in overall performance between students in the HR class and the CCC class on the pretest, there was a significant difference in performance on the posttest [F (1, 2) = 4.64, p = 0.04). Further analysis delineated the extent of the differences in students' knowledge of astronomy concepts at the end of the course. Students who participated in the CCC course demonstrated greater understanding of orbits on the posttest (see Table 2). There was no significant difference in performance on items related to eclipses, seasons, phases, or timekeeping for students enrolled in either the HR or CCC class.

Table 2. *Variance Between Groups by Astronomy Concept*

	t-test*			Posttest Mean	
	t	P-value	t crit	HR	CCC
Eclipses	0.34	0.74	2.07	6.4	6.7
Seasons	1.96	0.06	2.07	6.6	8.5
Phases	1.18	0.25	2.07	4.0	4.8
Orbits	2.19	0.04	2.07	7.5	9.2
Timekeeping	0.30	0.76	2.07	2.5	2.5

* 2 tailed

Discussion

One of the primary goals of this research was to determine if the use of modeling-based inquiry can enhance students' understanding of astronomy concepts. Students in both the HR class and the CCC class showed improvement on their posttests after demonstrating limited knowledge of astronomy concepts at the beginning of the course. This indicates that using modeling-based inquiry as a pedagogical strategy can enhance students understanding of astronomy concepts.

The second goal of this research was to determine if different conditions of modeling-based inquiry (HR and CCC) have different influences on students' astronomy knowledge. This was indeed the case as only students in the CCC class showed a statistically significant improvement in their overall astronomy knowledge. Likewise, although students' knowledge of concepts related to eclipses, seasons, phases, orbits and timekeeping were similar at the beginning of the course, at the end of the course, students in the CCC class performed significantly better on posttest items related to orbits than did students in the HR class.

There are a few likely causes that can explain the difference in students' performance on overall measures of astronomy knowledge. It may be the case because the CCC course is focused solely on developing understanding of astronomical phenomena from the current scientifically accepted perspective, students performed better than the students who participated in the HR course, where students have to move from rejected perspectives to the currently accepted perspective. This was reflected in the CCC students' performance on the posttest that focused solely on measuring knowledge of astronomy concepts associated with a heliocentric understanding of the solar system (the currently accepted perspective). Another factor that may have influenced the difference in students' performance is the fact that the spring 2003 semester was the first time the HR curriculum had been implemented. In contrast, the CCC curriculum had been used in this course for over 2 years prior to the spring 2003 semester, and thus there had been ample time to sort out issues related to effective implementation of the curriculum.

The fundamental differences between the HR and CCC conditions also likely contributed to the performance differences on items related specifically to orbits. The orbital relationship may be particularly confusing to students in the HR class because they are required to move from a geocentric model (the sun and moon orbiting around the earth) to a heliocentric model (earth orbiting around the sun). Students in the CCC class on the other hand, focus solely on understanding the heliocentric model of the solar system.

Conclusions

This research has shown that based on a quantitative assessment of students knowledge of astronomy concepts, using a modeling-based inquiry approach can be an effective method for enhancing students understanding of astronomy. Results also suggest that Constrained Conceptual Coherence and Historical Recapitulation approaches used within modeling-based inquiry pedagogy differently influence students' overall knowledge of astronomy. Students who participated in the class using the Constrained Conceptual Coherence approach performed significantly better overall, as well as on items specifically related to orbits, on the posttest.

We recognize that objective measures such as multiple-choice tests may not be sensitive enough to provide a complete or accurate picture of students evolving knowledge of astronomy concepts. Future research should examine issues and implications related to transfer of knowledge gained from a course that uses modeling-based inquiry approaches to traditional assessment measures such as multiple-choice tests. Future research should also be done to further examine how different conditions such as Historical Recapitulation and Constrained Conceptual Coherence influence student understanding of astronomy concepts when used within a modeling-based inquiry approach.

References

Marschall, L. A. (2000). The Universe on a Desktop: Observational Astronomy Simulations in the Instructional Laboratory, *Publications of the Astronomical Society of Australia, 17* (2), 129-132.

Prather, E. E., Slater, T. F., and Offerdahl, E. G. (2002) Hints of a Fundamental Misconception in Cosmology *Astronomy Education Review, 1*(2), 28-34.

Skala, C., Slater, T. F., and Adams, J. P. (2000). Qualitative Analysis of Collaborative Learning Groups in Large-Enrollment Introductory Astronomy, *Publications of the Astronomical Society of Australia, 17* (2), 185-193.

Sadler, P.M. (1992). Project STAR - Astronomy Concept Inventory. Harvard University

Wandersee, J. H., Mintzes, J. J., & Novak, J. D. (1990). Research on alternative conceptions in science. In D. L. Gabel (Ed.), *Handbook on Science Teaching and Learning* (pp. 177-210). New York: Macmillan Publishing Co.

Zeilik, M., Schau, C., Mattern, N. (1998). Misconceptions and their change in university-level astronomy courses. *The Physics Teacher, 36* (2), 104-107

The Effects of Source of Motivation and Mastery Contigency on Academic Achievement and Attitudes During Computer-Based Learning

Jale Balaban-Sali
Anadolu University
Eskisehir, Turkey

Abstract

This presentation involves a study in which it was examined to find out the effects of sources of motivation and mastery contingency on student's achievement, attitudes, time-on-task, confidence, retention, and motivation level. The sample of the study consisted of 137 sixth and seventh grade students. After developing two different computer-based instruction materials, the students were asked to study these materials and take the achievement tests, and attitude scale, and instructional materials motivation survey. One of the main results of the study was that the mastery level had a significant effect on academic achievement.

Background

Motivation to learn has long been considered an important concept of educational psychology and has been viewed by the major predictors of academic achievement and work productivity along with ability, instruction, and feedback. Various studies have indicated that motivation accounts for as much as 38% of the variance in student achievement (Means, Jonassen & Dwyer, 1997). Therefore, understanding of motivation and how they apply to instruction is important for teachers and instructional designers.

Also, motivational characteristic of the computer has been considered as an important factor in many computer-based instructional programs. Intrinsic features of the computer such as immediate feedback, animation, interaction, and individualization are more likely to motivate students to learn than any other media.

Keller (1983) has proposed the model of motivational analysis for an effective design of instructional strategies. Keller's ARCS model of motivational design introduces four basic categories for motivational tactics. Attention to capture the interest of learners and stimulate curiosity to learn, relevance to meet the personal needs and goals of the learners, confidence to help the learners develop positive expectancies for success, and satisfaction refers to positive feelings about one's accomplishments and learning experiences. ARCS model provides strategies that a course designer or teacher can use to make instruction responsive to the interest, curiosity, and needs of the learners.

The theory of mastery learning is based on the simple belief that all children can learn when provided with conditions that are appropriate for their learning. Mastery learning is essentially an instructional technique for the teaching and learning of hierarchical and sequential material. Material to be learned is subdivided into units or steps, covering from one lesson to several weeks' lessons. Students are given a test at the end of a unit, and if they do not achieve a mastery grade on the test (typically 80-95 %) they are provided with more time and more teaching until they can achieve a mastery grade on a retest (Arlin, 1984). That is, students who do not achieve mastery are remediated through tutoring, peer monitoring, small group discussions, or additional homework. Additional time for learning is prescribed for those requiring remediation. Students continue the cycle of studying and testing until mastery is met. Bloom (1976) suggests that at the end of an instructional unit, the teacher gives a formative (not used for grading) test to find out what has been learned and not been learned; determines corrective instruction for common errors, reteachs, perhaps in a different way or style, and tests again on the same items using altered questions. Grading for mastery in not on the curve. It means that every student can get an "A" if they master the material.

Bloom (1976) denies that some students are "good learners" while others are "poor learners." Instead he says that most students become very similar with regard to learning ability, rate of learning and motivation for further learning when they are provided with favorable conditions.

The primary purpose of this study was to determine the effects of sources of motivation and mastery contingency on student's achievement, attitudes, time-on-task, confidence, retention, and motivation level. In order to achieve this purpose, the following steps were followed: (1) Identification of audience characteristics; (2) conducting instructional analysis; (3) audience motivational analysis; (4) selecting motivational strategies; (5) development of computer-based instruction lesson; (6) conducting formative evaluation; (7) experimental study.

Method

Design and Participants

A 2 x 3 factorial design was used for this study. The independent variables in this study were the sources of motivation (intrinsic motivation and extrinsic motivation) and mastery contingencies (50%, 70% and

90%). The dependent variables were student's achievement, attitudes, time-on-task, confidence, retention, and motivation level.

The sample of the study consisted of 137 sixth and seventh grade elementary students (45% girls and 55% boys) in two private schools in Eskisehir. These schools were selected because computer was available for students use. Computer laboratories were also readily available in each of the two schools in the study. All participants were randomly assigned to one of six conditions of two independent variables.

Instructional Materials

Two versions of a computer-based instructional program on earthquake were designed and developed. These were intrinsic motivation version and extrinsic motivation version. Encouraging, motivational, and explanatory feedback were given to students in intrinsic motivation groups. Corrective feedback and an extrinsic reward (chocolate) were given to students in extrinsic motivation version. Explanatory feedback provides the learner with additional information. That is how or why a response is correct or incorrect and which pages of the instructional program were to correct it. Corrective feedback provides to the learner the correctness or incorrectness of a response such as "Sorry, your answer is not true". Encouraging, and motivational feedback consist of some words and phrases about learners progressing such as "That's amazing, you've mastered that, marvelous, nice going, good work". Besides, each of the two versions of the instructional program was developed three different criterions of mastery level (50%, 70% and 90%). Thus, totally six versions of the program were developed. This program was developed using Macromedia Director 8.0.

Each of the computer-based instruction programs was divided into four modules, each with their own objectives and assessment. Students had prerequisite skills to move to next unit. That is, student was to proceed to new unit until basic prerequisite unit was mastered. Each student would master at least 50%, 70% and 90% proficiency level of the material before proceeding. Students moved at their own paces.

The titles of the modules were "Why do earthquakes happen?", "What should I do before an earthquake?", "What should I do during an earthquake?" and "What should I do after an earthquake?" The content was adopted and modified from print-based and web-based earthquake resources. None of the students took earthquake content in their science lessons before. Therefore, studying instructional materials on the earthquake was expected to be more relevant for students because of the big earthquake experiences in 1999, Turkey. The researcher believed that course relevant would exert a positive effect on learning, because as Dweck and Elliot (1983) observed, if students have a reason for engaging in a task the importance of the task is apparent, then academic achievement will be improved.

Selection of motivational strategies was conducted through the following third stages. First, potential strategies were obtained from the survey and interviews with science teacher and from the motivational strategies suggested by Keller and Suzuki (1988). Second, the criteria for selecting motivational strategies were identified. Keller (1987) suggested some guidelines for selecting motivational strategies. These suggestions were established as criteria to be used for selecting motivational strategies for design and development. Third, the appropriate motivational strategies which met the established criteria were selected from all the obtained motivational strategies. In this stage, motivational strategies divided into two categories: strategies for intrinsic motivation and strategies for extrinsic motivation.

The validation of computer-based instruction programs was done through a formative evaluation. Two expert instructional designers, two content expert about earthquake, and one elementary school science teacher participated in validating of the CBI. In the pilot study, 18 elementary students participated in the formative evaluation. Three students went through each version of the CBI. After they finished the program, they were given the achievement test, attitude scale and the IMMS. Students' feedback on the CBI programs was obtained.

The experiment procedure was similar for the six types of groups that were used. It was conducted at the two elementary schools in Eskisehir. Six 60-minute sessions were set up to observe the students. For each session, students came to the computer laboratory and were assigned to one of the six conditions by the researcher and one assistant. Students were allowed to study as long as they wanted and the amount of time spent each learner on computer-based learning was recorded by the computer. Students completed the computer-based material in approximately 1 hour. After they completed their learning, academic achievement, attitude, and motivation were measured.

Instruments

Following the instruction, all participants completed an achievement test, a Likert type attitude scale and instructional materials motivation survey (IMMS) originally developed by Keller. The researcher requested and obtained written permission from the author to use and/or modify the survey instrument for this study.

Achievement was measured by 30 multiple-choice test. The test was designed to measure the learners' understanding of the instructional content (capabilities included facts, concepts, and principles). The reliability of the achievement test was .70, using the K-R 20 formula. Same achievement test was given to students two weeks later after the experimental study to measure their retention levels.

Attitude was measured by 35 questions. Each questions was answered on a Likert type scale of 1 ("Strongly disagree") to 5 ("Strongly agree"). The Cronbach's alpha reliability of the Likert type attitude scale was .90. Motivation was assessed by the Instructional Materials Motivation Scale (IMMS) administered to the learners following the completion of the instruction. Keller (1993) designed this scale to assess the motivational characteristics of instructional materials. It consist of 36 Likert-scale items the measuring the four motivational levels (attention, relevance, confidence, and satisfaction) of the learners. Students were asked to respond on a five-point survey with response choices of 1 (not true), 2 (slightly true), 3 (moderately true), 4 (mostly true), and 5 (very true). The Cronbach's alpha reliability of the instructional materials motivation scale was calculated as .90.

Data Analysis

To test for differences in sources of motivation and mastery contingency among learners in the six versions of computer-based instruction, two-way analysis of variance (MANOVA) was used. MANOVA was followed by Fisher's LSD pairwise comparison procedure if a significant difference was found among groups. The relationship was regarded as statistically significant when the P value was <0.05.

Results and Discussion

Achievement

Table 1 shows that the source of motivation did not have statistically significant effects on academic achievement, but the mastery level had a significant effect on achievement. Ritchie and Thorkildsen's (1994) research results provide some evidence that knowledge of being in a mastery adaptive program contributes to increased achievement. The findings from this current study are not sufficient to provide an explanation of mastery learning programs increasing achievement. This may be explained with awareness, expectancy, and pay attention to lesson. This mastery criteria knowledge may alter their perception of control over the learning environment. A second possible reason for increased scores that much of extra time was spent in remediation and plus tutorial time in mastery learning programs. That is, mastery learning programs provide additional learning time for students. For this reason, academic achievement also may be part of the remedy of instruction.

Summing up, 90% mastery contingency group's students were more successful because of increase in active learning time to learn of the content. As a result, if mastery criterion would be high, academic achievement will increase. Both of the meta-analyses of research on mastery learning report positive effects of this method on student achievement (Guskey & Gates, 1986; Kulik, Kulik, & Bangert-Drowns, 1990). On a retention test taken two weeks after the study, %90 mastery students group retained more than others. But this was not a significant difference for the groups at the .05 level.

Table 1. *Two-way ANOVA Results for Achievement Scores*

	SS	df	MS	F	p
Sources of Motivation	17,35	1	17,35	.30	.863
Mastery Contingency	88,68	2	44,34	3,49	.033*
Motivation*Mastery	11,28	2	5,64	1,09	.338
Within Groups	1665,19	131	12,71		
Total	1780,4	136			

Motivation

As expected, the students for whom the study of the programs was consist of motivational and explanatory feedback more motivating than the learners for whom the instruction was consist of corrective feedback. The source of motivation had a significant effect on motivation level, [$F(1, 131) = 11.430$, $p <. 001$]. Students in intrinsic motivation groups showed higher points of ARCS categories than those in extrinsic motivation. Table 2 shows that two-way ANOVA results for motivation scores.

Table 2. *Two-way ANOVA Results for Motivation Scores*

	SS	df	MS	F	p
Sources of Motivation	4271,91	1	4271,91	11,430	.001**
Mastery Contingency	1392,33	2	696,17	1,863	.159
Motivation*Mastery	1299,67	2	649,83	1,739	.180
Within Groups	48961,31	131	373,75		
Total	56183,11	136			

Regarding attention, two-way ANOVA revealed a significant difference for the groups, [F (1, 131)= 12.255, p < .001]. Table 3 shows that two-way ANOVA results for attention scores. Naime-Diefenbach's (1991) research results also showed that students who completed the revised lesson with enhanced attention had a higher score on the attention subscale than the control group. Table 3 shows that two-way ANOVA results for attention scores.

Table 3. *Two-way ANOVA Results for Attention Scores*

	SS	df	MS	F	p
Sources of Motivation	766,85	1	766,85	12,255	.001**
Mastery Contingency	130,19	2	65,09	1,040	.356
Motivation*Mastery	165,68	2	82,84	1,324	.270
Within Groups	8197,15	131	62,57		
Total	9304,55	136			

Concerning relevance, two-way ANOVA revealed a significant difference for the groups, [F (1, 131)= 7.939, p < .006]. The result of the Means, Jonassen and Dwyer's (1997) study suggest that relevance of instructional materials to particular target groups of learners should be considered when designing instruction. Students in intrinsic motivation groups showed higher points of relevance subscale than those in extrinsic motivation in the present study. Besides, research result shows that most of the motivational instructional strategies of the relevance component, as outlined in the ARCS model of motivational design were significantly effective in increasing students' perceptions of motivation within a videotaped instructional lesson (Nwagbara, 1993). Table 4 shows that two-way ANOVA results for relevance scores.

Table 4. *Two-way ANOVA Results for Relevance Scores*

	SS	df	MS	F	p
Sources of Motivation	312,08	1	312,08	7,939	.006**
Mastery Contingency	153,58	2	76,92	1,954	.146
Motivation*Mastery	89,31	2	44,66	1,136	.324
Within Groups	5149,58	131	39,31		
Total	5717,84	136			

With regard to confidence, finding that there were no significant differences among the two groups. But, there was a significant interaction between sources of motivation and mastery contingency groups, [F (2, 131)=3.99, p< .021]. This can be interpreted that condition of review opportunities could have resulted in non-significant differences in confidence among the two groups. Research results show that review opportunities are effective strategies for increasing students' confidence (Bickford,1989). Therefore, a possible explanation for this finding may be due to the review opportunities. Table 5 shows that two-way ANOVA results for confidence scores.

Table 5. *Two-way ANOVA Results for Confidence Scores*

	SS	df	MS	F	p
Sources of Motivation	76,61	1	76,61	3,241	.074
Mastery Contingency	64,08	2	32,04	1,355	.261
Motivation*Mastery	188,65	2	94,32	3,990	.021*
Within Groups	3096,71	131	23,64		
Total	3428,41	136			

Fisher's LSD pairwise comparison procedures revealed that students in the intrinsic motivation group showed higher confidence than others. Students in intrinsic motivation and 90% mastery contingency group showed higher points of confidence subscale than intrinsic motivation and 50% mastery contingency group. Table 6 shows that Fisher's LSD pairwise comparison results of confidence.

Table 6. *Fisher's LSD Pairwise Comparison Results of Confidence*

	i50	i70	i90	e50	e70	e90
i50	1.0000					
i70	0.0270*	1.0000				
i90	0.0080**	0.7070	1.0000			
e50	0.2200	0.2970	0.1450	1.0000		
e70	0.7810	0.0130*	0.0040**	0.1310	1.0000	
e90	0.4040	0.1650	0.0720	0.7060	0.2660	1.0000

* $p<.05$
** $p<.01$

Respecting satisfaction, two-way ANOVA revealed a significant difference for the groups, [$F (1, 131)= 8.893$, $p < .003$]. Students in intrinsic motivation groups showed higher points of satisfaction subscale than those in extrinsic motivation. Table 7 shows that two-way ANOVA results for satisfaction scores.

Table 7. *Two-way ANOVA Results for Satisfaction Scores*

	SS	df	MS	F	p
Sources of Motivation	172,3	1	172,3	8,893	.003**
Mastery Contingency	60,77	2	30,38	1,568	.212
Motivation*Mastery	12,35	2	6,18	.319	.728
Within Groups	2538,15	131	19,37		
Total	2793,88	136			

It demonstrates that different mastery criterion was affected students' academic achievement and motivation differently. The study demonstrates that it is feasible to design motivating instruction for a self-paced learning setting, such as mastery learning and computer based learning. The results of the study also indicate that instructional design, which is based on the ARCS model and mastery learning improved students' affective and cognitive performance and provided a useful guide for instructional designers who want develop computer based learning materials.

The study also suggests possible areas for future research. First, it is necessary to investigate how the application of ARCS and mastery contingency adaptive CBI will work for different ages of learners. In this study, elementary school students were the target audience. However, if the audience were either secondary students or adult learners, what might be the results of the same experiment? Which reward might be given secondary students and adult learners?

Second, the size and characteristics of the sample used in current study limit the generalisability of the results. The sample size is small even though the children came from private schools. Learner may be characteristically different from the public schools' students. Further efforts should focus on study with children in public schools to increase generalisability of the results.

Third, research is necessary to investigate the relationship among prior achievement, cognitive aptitude, learning style, locus of control, continuing motivation, and different content in motivationally enhanced CBI in comparison with mastery contingency adaptive CBI.

Fourth, the results from the present study provide some support that students' affective outcomes might improve as a result of using CBI, but further descriptive, correlational, and especially experimental research is needed to verify these results.

Fifth, longitudional studies are essential to adequately investigate how students' affective and cognitive outcomes change over time as a result of using CBI. Also, future studies need to specifically examine what is the cultural diversity of the learners. Any such quantitative, qualitative or integrated the methodology should include and be sensitive to different cultural values variables as well as motivational strategies that differ across cultures.

Sixth, instructional activities designed with ARCS model in affective domain, mastery contingency in cognitive domain have potential worthwhile in computer-based instruction, web-based instruction and distance education. This study will provide the foundation that will foster research efforts for designing and developing a more advanced motivationally and mastery contingency adaptive lessons in different media.

In conclusion, the results of this study show the possible value of the ARCS model as a guiding process for designing intrinsic motivationally enhanced computer-based learning. Further research should lead to more effective applications of motivationally enhanced design.

References

Arlin, M. & Webster, J. (1983). Time costs of mastery learning. Journal of Educational Psychology, 75(2), 187-195.

Bickford, N.L. (1989). The systematic application of principles of motivation to the design of printed instructional materials. Unpublished doctoral dissertation. Florida State Univeristy, Tallahassee.

Bloom, B.S. (1979). Human characteristics and school learning. NY: McGraw-Hill.

Dweck, C.S. & Elliot, E.S. (1983). Achievement motivation. In P.H. Mussen (Series Ed.) & E.M. Hetherington (Vol. Ed.), Handbook of child psychology (4th ed., pp.643-691). NY: Wiley.

Guskey, T.R. & Gates, S.L. (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classrooms. Educational Leadership, 43(8), 73-80.

Keller, J.M. (1983). Motivational design of instruction. In C.M. Riegeluth (Ed.), Instructional design theories and models (pp.383-434). Hillsdale, NJ: Erlbaum.

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. Journal of Instructional Development, 10(3), 2-10.

Keller, J. M. (1993). Instructional materials motivation scale (IMMS). Unpublished manuscript. Florida State University, Tallahassee.

Keller, J.M. & Suzuki, K. (1988). Use of the ARCS motivational model in courseware design. In D.H. Jonassen (Ed.), Instructional designs for microcomputer courseware (pp.401-434). Hillsdale, NJ: Erlbaum.

Kulik, C.L.C., Kulik, J.A. & Bangert-Drowns, R.L. (1990). Effectiveness of mastery learning programs: A meta-analysis. Review of Educational Research, 60(2), 265-299.

Means, T.B., Jonassen, D.H. & Dwyer, F.M. (1997). Enhancing relevance: Embedded ARCS strategies vs. purpose. Educational Technology Research and Development, 45(1), 5-17.

Naime-Diefenbach, B.N. (1991). Validation of attention and confidence as independent components of the ARCS motivational model. Unpublished doctoral dissertation. Florida State Univeristy, Tallahassee.

Nwagbara, C.I. (1993). Effects of the relevance component of the ARCS model of motivational design. Unpublished doctoral dissertation. Purdue University, West Lafayette.

Ritchie, D. & Thorkildsen, R. (1994). Effects of accountability on students' achievement in mastery learning. Journal of Educational Research, 88(2), 86-91.

This study was supported by Research Grant No.000159 from the Anadolu University, Turkey.

Correspondence concerning this study should be addressed to Jale Balaban-Sali, Faculty of Communication, Anadolu University, Turkey. Electronic mail may be sent to jbalaban@anadolu.edu.tr.

Progressive Evaluation in a Media Production Course

Doris U. Bolliger
St. Cloud State University

Abstract

This paper outlines the progressive evaluation of an instructor by students who taught a graduate level face-to-face, computer-based authoring course. Individuals enrolled in the course had the opportunity to evaluate the instructor using formative and summative evaluation methods. Students provided the instructor with feedback on a weekly basis using a three-minute paper and completed a course evaluation questionnaire at midpoint and at the end of the semester. The feedback was used to provide learners with a review and clarification process and to revise the course on an ongoing basis. Results show the increase of ratings collected with the use of the course evaluation questionnaire was statistically significant at the .01 alpha level.

Introduction

Students typically complete course evaluation forms at the end of the semester in order to provide instructors or administrators with feedback pertaining to the instructor's teaching skills or administrative support services. Feedback from students gathered through the summative evaluation process is helpful in rethinking instructional methods and strategies, revising course materials and content, and in improving student services for the upcoming semester. However, when we, as instructors, do not use formative evaluation, we miss the opportunity to improve our methods or materials throughout the semester. We should, therefore, ask our students to provide us with feedback on an ongoing basis. Questions such as what was clear, interesting, and useful during each class session can provide much needed insight into our students' learning process.

Turning the Table

Instructors provide students with feedback on an ongoing basis throughout the semester. We design several assignments that are due at various points during the semester, give students the opportunity to revise projects, and grade their work. Students should have a good idea what their levels of performance are and what their current standings are at any given time. In other words, we provide them with formative *and* summative evaluation. Should we not give students the opportunity to provide instructors with the same type of evaluative process? After all, instructors should strive for continued improvement in their teaching.

When we employ both formative and summative evaluation methods, we communicate an educational philosophy to our students. We make them aware of the fact that we value their opinions and thoughts, and we can respond to them in a timely fashion. Feedback should be of benefit to individuals who have taken the time to provide it, rather than to the persons who will be in the course the next time around. In addition, this process can be perceived as a value added benefit to our learners.

Background Information

The researcher decided to implement formative evaluation methods in addition to the summative evaluation method during the spring 2003 in a graduate level computer-based authoring course required for a Master of Science Degree in Instructional Design and Training; one of the course prerequisites was the successful completion of a graduate-level instructional design course. The course was taught in a face-to-face classroom environment; however, students used Web-based tools extensively. Course materials such as a syllabus, assignments, a schedule, and resources were posted on a course Web site and in WebCT, a Web-based course management system. Additional materials such as PowerPoint presentations and threaded discussion questions were posted in WebCT.

The instructor desired to detect any stumbling blocks in the learning process early on and provide assistance to students who needed additional help. Students were required to learn several new software programs in order to complete their projects. All multimedia projects needed to be uploaded to their Web sites on the university server. Writing assignments were submitted via e-mail or with the use of the WebCT assignment tool. The researcher assumed students had prior experience using WebCT and some of the authoring software used before they entered the course.

Research Methodology

Data were collected with the use of two instruments. A modified version of the *Minute Paper* (Angelo & Cross, 1993) was used in this study. During the course of the semester, students were encouraged to complete a paper-based questionnaire on a voluntary and confidential basis every week after the class session concluded, with the exception of the last session. This questionnaire, the *three-minute paper*, had three 4-point Likert-type questions ranging from 1: strongly disagree to 4: strongly agree. Students were asked to indicate their level of agreement with statements pertaining to clarity, interest, and usefulness of the session. In addition, students were asked to fill in responses to three open-ended questions relating to the most important thing they had learned that day during class, questions that remained unanswered, and how the session could have been improved.

The researcher hypothesized students would be able to provide the instructor with valuable information during the formative evaluation process which would assist in the dynamic revision of the ongoing course. The instructor expected she would be able to use weekly feedback to review material that was not comprehended by students in the previous week. The researcher also expected several suggestions on how to improve some of the class sessions.

Students also completed a paper-based evaluation form at midsemester and at the end of the semester; participants were assured anonymity and confidentiality. This student satisfaction questionnaire had 25 four-point Likert-type questions pertaining to the instructor, technology resources, administration, and course materials. Two open-ended questions asked students about the least and most satisfying aspect of the course. The questionnaire is a modified student satisfaction survey that was developed by Bolliger and Martindale (in press) for the evaluation of online courses. Student ratings on the course evaluation were not expected to change significantly.

The collected data was analyzed using quantitative and qualitative methods. Missing data were replaced with the mean and data was examined for univariate outliers. A *t* test was performed in order to determine if ratings were statistically significantly different on the mid- and final-course evaluation. The reliability of the instruments was determined calculating the Cronbach alpha. The qualitative data was analyzed using open coding to create categories.

Results

Weekly feedback. On the three-minute paper, at least 87.5% of participants strongly agreed or agreed that each class session was interesting and useful. When asked if each session was clear, at least 67% strongly agreed or agreed that it was. The highest and lowest mean for each of the three variables was as follows: 3.75 and 2.89 for clear; 3.57 and 3.00 for interesting; and 3.71 and 3.00 for useful, respectively. Figure 1 displays the means for the three variables for each session. The means for the variables fluctuated between weeks and did not always correlate within sessions.

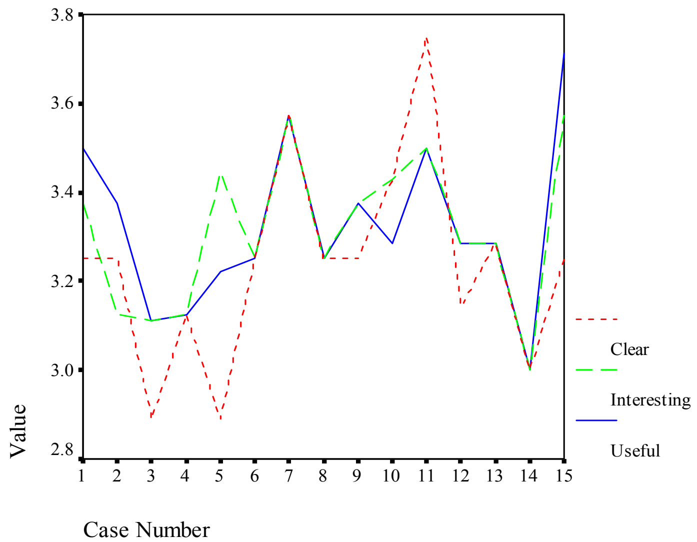


Figure 1: Variables and ratings for each of the weekly class sessions.

Even though the quantitative results were important to the instructor, the most valuable were responses to the open-ended questions on the survey. Students provided several questions that remained unanswered and provided perceptive feedback to important things learned and suggestion for improvement. These comments were analyzed and discussed with the learners at the beginning of each following week's session.

As expected, the number of comments pertaining to important things learned and unanswered questions remained stable over the course of the semester. However, the number of suggestion that were listed by students

dropped as the semester continued. Particularly during the first three weeks of the semester students were unsure about course requirements and needed clarification, even though most of the information was provided on the course Web site.
Most of the items about important things learned related to content such as instructional design elements; software use; project ideas and clarification; review of previously covered content; writing; and completed projects of peers. Many entries indicated that no questions remained unanswered. Other comments addressed software use, and course management issues such as workload and grading. Several answers to the question pertaining to improvements revealed that students did not have any suggestions. However, several suggestions were made referring to slowing down the pace, bringing in additional examples, increasing interaction between peers, providing more details on the projects, and reducing readings for weeks students wanted to concentrate on software applications.

Midsemester feedback. After the course evaluation was completed it became clear that a few students did not like the threaded discussions ($M = 2.38$). They disagreed somewhat that the content was well organized (M=2.88), the textbooks were useful ($M = 2.71$), and the workload was too much ($M = 2.38$). They felt that posting to threaded discussions and generating exam questions was "busy work," and asked for additional course time to practice with software programs and work on their group project.

After reviewing the feedback provided by students, the instructor reduced the number of threaded discussions and number of exam questions which had to be generated. Students were given more time to work on their projects during class time; however, that would have been the case regardless of their suggestions. It was the nature of the course to introduce most concepts and theory related topics during the first half of the semester and provide time for group work and hands-on development during the second part of the course.

Final-course evaluation. The majority of students agreed or strongly agreed with all 25 items on the final course evaluation. Only four items received a mean below 3.0. After the semester was over, some respondents were not satisfied with the access to course materials (M = 2.89). Others still did not think the textbooks were useful ($M = 2.89$); they did not like the workload ($M = 2.67$), nor did they like the threaded discussions (M=2.78). Students explained they would like to have printed copies of articles. They thought the cost of textbooks was too high, and they did not like "doing busy work in a graduate-level course." Even though four items each received a mean below 3.0 on the mid- and final-course evaluation, students gave higher ratings on the final course evaluation questionnaire. The increase in ratings was statistically significant ($p = .01$). The null hypothesis was therefore rejected.

Both instruments used in this research study had acceptable reliability coefficients. The overall internal reliability of the three-minute paper was high (.90). One hundred and thirteen surveys completed over a 15-week period were used in the reliability analysis. The reliability of the original student satisfaction survey (Bolliger & Martindale, in press) had been established previously; its Cronbach alpha was .97. The modified version of the student satisfaction questionnaire had a high reliability coefficient (.91).

Conclusion

One issue that should be addressed here is the fact that the change in course ratings was statistically significant. Lewis (2001) pointed out similar findings. Seeking feedback from students and addressing their concerns and suggestions communicates the instructor's willingness to listen to students and commitment to student learning. Tenure-track faculty members are typically evaluated on the basis of student course ratings. If other researchers can validate the findings of this study, it may be advantageous for faculty members to implement the administration of midsemester course evaluations.
However, readers are cautioned not to generalize these findings to other populations. The sample in this study was drawn from only one geographical region and was very small ($N = 9$). Further replication studies need to be conducted in order to validate these results, particularly with a larger sample. The researcher plans to continue this research and to use formative and summative evaluation methods in this instructional media production course in the future.

References

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.).
San Francisco, CA: Jossey-Bass.

Bolliger, D., & Martindale, T. (in press). Student satisfaction in an online master's degree program in instructional
technology. *International Journal on E-Learning.*

Lewis, K. G. (Ed.). (2001). *Techniques and strategies for interpreting student evaluations.* San Francisco, CA: Jossey-Bass.

Using Social, Self-Directed Learning Frameworks to Engage and Transform Aspiring School Leaders

Naomi R. Boyer
University of South Florida

Abstract

Life-long learning, reflection and teamwork are all attributes that enhance leadership capacity and school reform efforts. Yet, the skills and the process to build these skills are not traditionally integrated into university coursework in teacher education or aspiring leader programs. A framework has been developed that incorporates social experiences, self-direction, metacognition and learning engagement into a technology integration course. This framework includes a series of scaffolds that were instituted in five successive semesters of the identical web-based course. A global model of social, self-direction has been developed to provide a theoretical foundation to support the process of this framework and is being studied through the lens of design-based research methods. Narrative accounts, surveys, and reflective instruments were used to garner data about student satisfaction and learning progress in investigating the question: What is the impact of social, self-direction scaffolds on aspiring school leaders' ability to plan, manage, sustain and complete personal/group learning experiences? These scaffolds have been found to have a tremendous impact on aspiring leaders' ability to plan, manage, sustain and complete self/group design learning experiences. Students report a transformation in perspective as well as ability and report large scale impact upon authentic school environments. Further investigation would need to be conducted to evaluate long-term influence and additional model iteration.

Introduction

The online environment provides an opportunity to reform current educational practice by adapting teaching pedagogies to encompass adult learning theory. It is imperative in higher education that adult learners be granted the freedom and the guidance to actively engage in the learning process at both the undergraduate and the graduate level. The "Seven Principles for Good Practice in Undergraduate Education" provided by Chickering and Gamson (1987) can be used to gauge effective learning in both face to face and distance opportunities. These principles include: student/faculty contact, active learning techniques, prompt feedback, cooperation among students, time on task, and communication of high expectation (Skill & Young, 2002; Chizmar & Walbert, 1999). What emerges at the juncture of "good practice" and adult learning "self-directed" philosophies is a framework that encourages personal responsibility of learning, social engagement, reflective practice and continued life-long learning strategies.

The framework and model of learning described above will be introduced throughout this paper. Issues such as reflective practice, learning scaffolds, self-directed learning, and learning patterns will be reviewed. This study explores the following basic research question. What is the impact of social, self-direction scaffolds on aspiring school leaders' ability to plan, manage, sustain and complete personal/group learning experiences? A model of social, self-direction was utilized as the program structure and the research method includes the use of design-based research approaches for investigating effectiveness. A study has been conducted with aspiring school leaders and has been through five iterations of adaptation. The findings will be described and conclusions explored.

Designing Educational Environments That Support Leadership Development

Leadership development, particularly school leadership, has been found to be closely tied to self-knowledge which can be facilitated via personal reflection (Kouzes & Posner, 1995). Barth (2001) imparts the importance of reflection in the following excerpt:

It is through reflection that we distill, clarify, and articulate our craft knowledge….Reflection is precisely the capacity to distance oneself from the highly routinized, depleting, sometimes meaningless activities in which we are engaged, so that we can see what's really going on. (p. 65) School leaders also need to engage in goal setting to demonstrate and model for their organizations the way to continue learning and personal improvement (Barth, 2001). Including reflective opportunities and reflective devices throughout the process of learning is a way to regularly think about leading and learning (Lambert, 2003).

Using self-directed learning scaffolds in online environments is a highly effective way to both involve students in the act of learning and re-direct the traditional framework from "teacher-centered" to "learner-centered" instruction. The establishment of scaffolds to support novice learners coupled with a steady progression to remove structure as students gain knowledge, confidence, and skills can increase independence and encourage self-regulated patterns (Dabbagh, 2003). The term scaffold here is used to represent the creation of a support system to assist learners in the development of a personal learning plan coupled with reflective mechanisms that provide a clear structure, yet allow for the removal of traditional boundaries as the learning process progresses. Some noted boundaries that can be minimized through the use of these tools include: instructor identified final products, non-relevant, meaningless work, date/time limitations, lack of flexibility and adaptation, and narrow confines of evaluation and assessment.

Self-directed learning has generally been considered an individual facet of the learner rather than oriented toward community learning experiences (Long, 2000). Consequently the use of self-directed tools such as learning contracts, diagnostic instruments, and reflective participation have not been explored for application in joint endeavors, as a means to identify a common goal, plan, and overall educational direction. Students using these frameworks have the ability to invest in the learning process, utilize meta-cognition to relate to others and participate in the assessment and evaluation to reflect on learning outcomes. Combining the concepts of socio-cultural learning (Vygotsky, 1978) with the adult learning philosophies of self-direction (Knowles, 1975) provides a landscape for the emergence of an active learning model that minimize isolation, propels learning community development, and institutes a means for communication, planning and individualization.

In a model developed by Boyer (2000), the online learning environment is viewed as a complex system. The initial model has been re-designed to accommodate the aspects of online, social, self-direction included in this study (see Figure 1). The model utilizes a systems framework of input-process-output that is surrounded by a continual feedback loop. The specific question of interest in this study was designed to investigate the appropriateness of self-directed instruction with aspiring school leaders, which relates directly to the process portion of the model.

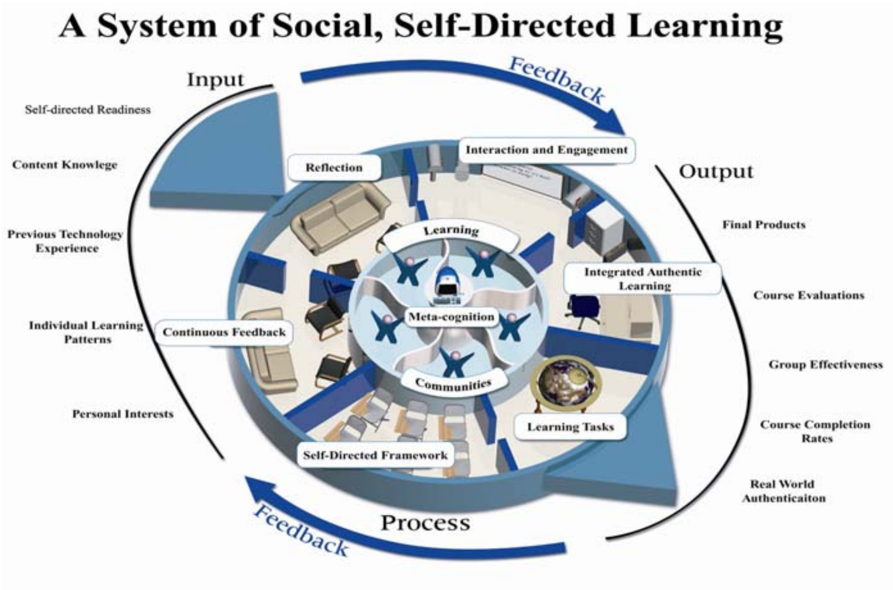


Figure 1. A systems model of social, self-directed learning incorporating the online environment into the process elements.

Methods

The Implementation Design of the Scaffolds

A course design which incorporates social, self directed learning scaffolds, reflection (meta-cognition), leadership development philosophies and "good practice" has been adapted within five successive semesters (20 month time frame) of the identical course title. Two additional sections are currently in session. This study involved 123 Educational Leadership graduate students in a technology integration course required within the Masters program. As a web-based designated course, students spent more than 75% of class activity online. Essentially, the course is blended in format with an initial course orientation meeting (7 hours) and a final sharing session (3 hours). All other course interaction and work is completed online. The course structure was modified based on student feedback, results of the Learning Combination Inventory-LCI (a cognitive learning styles instrument), course objectives, diagnostic instruments, learning contracts, online discussions, and course products. A course evaluation was also designed to gather information on particular facets of online course satisfaction.

Each of the data collection tools mentioned above was also utilized as a learning strategy for student development. For instance, the LCI is given to students at the beginning of the semester to not only collect valuable group information, but also to provide self-awareness and strategies that assist individuals in planning and completing individual and group work. This learning pattern information is then used strategically throughout the semester to have students meta-cognitively examine their learning processes as they progress individually and collectively throughout the course. Group compilations of the LCI are shared a few weeks into the semester so that all have an awareness of the group's strengths and weaknesses and then helpful hints are given to make their experience more successful given the make-up of the group. A researcher who is part of the research team provides this information to the groups and maintains a role of "meta-cognitive coach" throughout the semester to assist groups in processing learning pattern information.

A learning contract process is provided and clearly defined as the structure for students to guide their online learning experience. Students begin by completing a diagnostic instrument in which they rate their current knowledge base and necessary proficiency for professional competence on course objectives. This diagnostic instrument then provides the information necessary for creating individual and joint learning contracts. The group (or individual) then selects five objectives that they are less knowledgeable about and that they value as important to their role as a future administrator. Of the five objectives, at least three objectives are required as part of the group work, while two would be completed individually. Groups can choose to complete all five objectives as joint projects. These objectives are then transferred to the learning contract, strategies and resources are identified from course content and other materials, dates are selected, evidential products planned for final objective demonstration and decisions made about who and how the final evidence will be authenticated (verified for quality and content). Students can request workshop meetings, additional materials, and other face-to-face experiences as part of the resources and strategies. At the end of the semester, students complete a grading contract to reflect on their individual/group competency and completion of learning contract objectives. The steps of this process are explained, demonstrated, and practiced during the full-day session "orientation" at the outset of the course.

Students are required to participate in the online setting to maintain an engagement in the course and a connection to others. Participation is defined in the class as logging on the course site at least three times a week, answering e-mails in a timely basis, and reading discussion room threads and responding where appropriate. Students evaluate/reflect on their own participation rates, quality and quantity through an online-discussion self-rating form, which is included in the final grading procedures. All technologies are encouraged to facilitate group member contact, such as chat, instant messaging, video-conferencing, conference calls, faxing, and face-to-face meetings when necessary. Each student is also required to complete a student homepage to facilitate classmate recognition and networking aspects of learning. Feedback is provided to all students on a daily basis, via discussion board and e-mail responses.

Not only is the scaffolding process aligned to what is known about learning and retention found in current research, but this process also leaves an audit trail that is clearly helpful in most university accreditation and record keeping processes. The continual improvement, student feedback, and work outcomes establish evidence of successful learning while also satisfying guidelines for quality control assurances. This scaffolding also increases reflective activity needed to integrate self-awareness and greater content understanding.

Data Collection

Design-based research methods were utilized to further involve theory development, model replication and empirical emphasis in the investigation of the innovated scaffolds for facilitating the social, self-directed phenomena. The design-based research format incorporates collaboration, model verification, complex system investigation, diverse data collection techniques and reform (Cobb, Confrey, diSessa, Lehrer & Schauble, 2003; Shavelson, Phillips & Feuer, 2003, Sloane & Gorard, 2003). This particular study question (What is the impact of social, self-direction scaffolds on aspiring school leaders' ability to plan, manage, sustain and complete personal/group learning experiences?) was used to explore the outcomes of the systems model on the aspiring leaders studied in this implementation of the design. Outcomes here refer to the perceptions of personal relevance, impact on external organizational environments and improved learning self-efficacy. Despite the overall continuing research using design-based research methods to improve and refine model design during iteration testing, this particular study incorporated open-ended narrative questions and surveys to specifically study the phenomena in question.

Students were asked to complete periodic personal reflections that were initially offered at the outset and the finality of the program and then in later phases emerged as a reflective tool throughout the course (approximately 5 times throughout a semester). Some of these reflections were offered via email, while others were provided on the online discussion board open to all students. Students always maintained the option of sending reflective comments directly to the professor via email, to ensure privacy issues. The grading contracts and online rating forms mentioned as part of the implementation design also contained valuable personal information as to the student's perception of the experience and were submitted privately to the professor.

Researcher designed course evaluations also provided valuable completion information including course relevancy, significant student learning, perspective shift and overall satisfaction. These evaluations included 18 questions and were submitted anonymously to the professor during the last week of class and were in no way linked to student grades. Students could choose to opt out of completing these non-university sponsored evaluations.

These tools were compiled across semesters and analyzed for general themes that emerged across reflections, forms and surveys. Three semesters of reflections and five semesters of course evaluations were used as substantive data to determine student personal growth and transformation. Given the design-based approach used in the comprehensive research, each semester contained iterations to address issues that were raised via the reflective tools mentioned above (Cobb et al., 2003). Therefore, while the overall structure of the course was identical in four out of five semester offerings of this class (the first pilot semester was significantly different), some interventions and curricular alterations were instituted to further enhance educational outcomes. The research methods described above were embedded into the learning features of the course rather than as external instruments, thereby reducing additional student work load and increasing student commitment. Descriptive statistics were compiled from these sources and were used to support themes.

The roles of researcher and professor where held by the same individual, however a research team was continually involved through collaboration to design interventions, supply content for students, and collect and analyze data, which is described by Zaritsky, Kelly, Flowers, Rogers & O'Neill (2003) as integral to the design-based approach. This research was conducted in an authentic setting and includes a variety of mixed methods that are tied in this phase of model-design research to narrative accounts.

Results

All 123 participants in the study (all of those enrolled over the five successive semesters) successfully completed the course. This high completion rate indicates a sustainable online learning environment with no student attrition despite a student population with diverse technology skills. Every student in semesters 2-5 completed the following tools: diagnostic instrument (self-rating), learning contract, grading contract and online participation rating form. Learning contract objectives were individualized to group and personal needs and were achieved at a variety of competency levels; however all participants earned letter grades of B or higher. The initial orientation session, the full day overview of course materials, online courseware, learning style exploration and learning contract theory/application, was found to be overwhelming to the majority of students. However, the disequilibrium that was established at the outset was greatly reduced as the learning contract planning process was complete. One student, Kathy indicates, "I have learned today that I feel very uncomfortable when I don't know EXACTLY what is going on. I think once I actually get started on creating and finishing our goals I will feel much better". This sentiment was reinforced throughout the semesters, and while students felt off-balance, an atmosphere of continual feedback, mutual trust and supported risk taking was established to assist in the transformative learning process. The overwhelming nature of this initial face-to-face meeting was explained to students as part of the natural learning curve and was reinforced by an e-mail from the professor to all students. Betty, a student in semester 4, responded to the orientation reflection prompt and professor assurance, "Okay! I thought I was the only one with that feeling, a sort of internal warning system letting me know I was in the wrong degree program! But of course, with a little time of "reflection", I felt much better about it. So thanks for letting me know I was NORMAL!"

Students had time during the initial session to collaborate in their self-organized group to establish a first draft and/or direction for their joint objectives (as previously mentioned students could choose to do 3-5 group objectives) with professor support and guidance. These draft documents were then electronically submitted after the first week of class. Professor feedback and suggestions were then sent back to the student groups (same process employed on individual objectives). Students made negotiated changes and then submitted final copy to the professor. Anxiety levels began to decrease as these learning contract documents were approved and learning planning complete. Even though personal learning autonomy was new for the majority of the students, by the end of each semester students reported course relevance on an open-ended course evaluation question as the following:

- I feel that the independent structure and accountability lends to optimum learning.
- I liked that I was able to learn what I wanted to learn not what someone told me I was going to learn. I got to choose what I wanted and that was great because I know what I am weak in and where I need help.
- I always wanted to take and online class. It [the course] definitely put me into a comfort zone and forced me to think about learning styles, group vs. individual work- it made me think.
- I really enjoyed doing activities that would benefit me! Many of my courses have been filled with busy work that is not applicable to my pursuit of an administrative position. I know that the skills learned in this course are relevant to me not only in my profession life, but my private life as well.
- It was practical and meaningful because I was able to choose areas that I felt were important. I chose the areas that I felt I needed to work on most. I realize now there is way too much to learn in one semester.

Therefore, despite an initial sense of fear, confusion and pressure directly derived from being overwhelmed by new skills, content and the unknown, the aspiring leaders in this program demonstrated an ability to plan their own learning environment. The reflective documents (online discussion rating forms and grading contracts) did indicate that some had difficulty maintaining appropriate time management skills and were "forced" to be conscious of time elements, in ways that other face-to-face courses did not require.

While all students met with a level of success, not all students enjoyed or realized the full potential of this format of learning. Each semester there were 2-3 students who reported having concern that without professor guided instruction (instructor lecture and assignment guided in a face-to-face environment) they were not garnering the necessary skills or knowledge. There was also a concern that if students did not expend effort that they could complete the course with little work. Mike suggested, "The way that the course is set up now ("pick-and-choose"), someone who is computer-savvy could join a group with a couple of people who doesn't even know how to turn on a computer and coast through the 10 weeks [summer session] without even working". This statement suggests a lack of personal accountability and a mistrust of the learning maturity level of others. While, this outcome is entirely possible within the scaffolds and design of the course, learners in any environment always have the option of "squeaking" buy with little to minimal work/commitment. Personal responsibility for learning is developed by engaged students who finds the value in continual growth, learning, and development despite the hard work and time investment.

The majority of student groups completed objectives that met a standard far exceeding instructor expectations. These Master's level students, in some cases, chose to conduct extensive research, collect data, analyze data and prepare presentations. Other groups completed presentations on legal, security and privacy issues, which were also delivered the local school district administration. Evidential products were planned, managed and completed that provided knowledge demonstration of both skill and conceptual competency. Many student groups also chose to integrate evidential products together reflecting a high degree of analysis and synthesis of initial objectives. In the final class reflection students reported:

- It's important for an educator and administrator to stay alert to new developments and always be willing to learn. I've learned that the Internet itself is an educator. There are many tutorial sites, information links and endless amounts of information. This class has also taught me that distance doesn't matter.
- There was a time when I would insist that I do most of the work, mainly because I didn't trust my partners to do a good job. Because of time constraints and knowledge about the personalities of my teammates, I was able to 'let go'. Learning about our strengths and being able to communicate to them was very powerful.
- I learned to be more independent with the computer. Just keep trying and eventually you will muttle your way through.
- I learned more than anything that learning is a great reward unto itself.

- I learned how to deal with my learning "discomfort" as well as my learning styles…So I learned a lot about myself and technology too.
- I look forward to a lifetime of using the tools that are available to me through Technology.

Students learned not only about the objectives that they selected, but about themselves; their strengths and weaknesses, their learning preferences as to face-to-face or web-base class opportunities and their ability to work with others to complete an end product were all indicated as side issues to the technology integration content of the class. A reduction of fear of the unknown and the ability to "play" with technology avenues to find solutions was also reported as significant learnings of the course. The end of the semester, half day meeting allowed students to see the final products of their classmates and groups and evaluate their accomplishments in terms of peer standards. No two group/individual projects, learning contract evidence and or presentations were the same and all marveled at the output projects that were accomplished in one semester's time, despite the differing learning levels. The learning experiences of the others provided materials and handouts to be included in personal materials to be used as a future administrators.

The impact of these scaffolds also proved to be extended into the lives of the students via school involvement and personal exchange. Student work was authenticated (quality and content verified) by experts embedded within their settings, who then provided documented feedback through a variety of methods. Many students had checklists, evaluation forms, written feedback, personal reflections and other forms of verification completed by technology specialists, principals, lead teachers, media specialists, district representatives and other experts "in the field". Students reported on final reflections and course evaluations, that course products and or new projects were being used in their current classrooms, school settings or districts. The time invested was having a direct impact on altering the authentic environments of the aspiring leaders, which would then provide opportunities for others to recognize their skill in preparation for new leadership positions.

Conclusions

At the conclusion each semester, it was evident that the self-directed learning framework was able to be successfully applied to a group learning format in the web-based setting. Rather than controlling the mechanisms of learning and disseminating knowledge, the online instructor was able to structure scaffolding, through which the facilitation of learning process was encouraged. The provision of opportunity was confirmed by student success, satisfaction and retention. It was possible to give attention to both group and individual requests throughout the course semester. At the end of each semester, many of the participant students felt that they were now better prepared to explore further learning in this format.

Learning contract documents supplied documentation of the aspiring leaders' ability to plan personal and group learning experiences to meet needs, interests and areas of importance. The learning contract included target date identification, which encouraged students to focus on personal time management and planning. Dates could be adapted as the learning contract represented a living breathing document; however, students were limited to the completion date of the class. Students reported that this was difficult at times, but that their personal planning skills improved due to this process. The final, successfully completed evidential products provide evidence of the sustainability, management, and follow through on work and plans. Involvement of external verification sources demonstrated student ability to expand learning into community activity and extend learning from a personal private enterprise into a socially constructed, authentic achievement.

Given the hands-on, learning focus of this approach and the skill based nature of the course content, student comments indicated an immediate transfer of knowledge to work environments. Students in both the classroom and administrative settings reported new technical skill usage. An example of this transference is expressed by Sally, "Even though my coursework is nearly finished, I have taken most of my learning goals and objectives and am now converting those for my classroom in August…..More work and this time no grade!" Semesters later students returned to report how they had used their new skills to complete school based tasks and or transform their schools in regard to technology integration. The value of this ongoing incorporation of skills within the work environment is meaningful for the student, the university and the school work place. Some also indicated an inclusion of a choice-based curriculum and/or learning pattern approach, in their own classrooms.

Future leaders require key abilities for successful leadership in schools and life. Some of these abilities include: reflection, personal responsibility, continual learning, time management, organizational facilitation and change, and strategic planning. Most traditional university courses do not provide strategies for addressing the characteristics listed above. The described course design has provided the opportunity for students to not only learn about the course content - technology integration, but also to use adaptable scaffolds with guided support and facilitation to take charge of personal learning experiences. Not all course participants reported a high satisfaction with this design and model. In fact, a minority reported preferring to be teacher guided with explicit instruction and assignments. There appeared to be a link between level of self-directed readiness and satisfaction within the social, self-directed environment. The students who needed to be other-guided rather than self-guided tended to express a need for additional face-to-face opportunities. Ongoing participation and engagement on the discussion boards was also a key element to those who reported success and increased knowledge.

Additional iterations to current semesters have been instituted to account for some of the issues that were raised by the groups described above. The basic course design has once again remained the same, however, weekly resources have been added as a prompt to address the personal accountability issue that was raised by a few concerned students. Many felt that there was so much information and material around them that they were unsure which was important to explore and which they should bypass. Therefore, these weekly resources provided quality materials that students could explore and discuss and could be used to further their contact objective completion. This appears to have been a highly successful intervention.

The particular study described above, focused upon a particular group of students in an isolated course. In this setting, the model has been quite successful and has been found to increase online engagement, reduce isolation and encourage the building of necessary leadership skills by having an impact on the personal and professional lives of the aspiring leader. Despite this success, the fact that this model is not replicated in other portions of the program limits the generalization of skills to other learning experiences. Statements such as, "I now have the hang of this, what other courses can I take that are of similar design" have frequently been heard. Other current research projects include the design of a program for health education leaders that applies this type of learning model to an entire program (a series of courses, in cohort format, that includes an overarching program contact). To continue the design-based research method and pinpoint additional portions of the model for study, other individuals, settings and studies will be added to existing research to extend the impact and implications into a larger population including more generalizable features and validity checking of constructs and schematics.

References

Barth, R.S. (2001). *Learning by heart*. San Francisco, CA: Jossey-Bass.

Boyer, N.R. (2000). *Building online learning: System insights into group learning in an international online environment.* Unpublished doctoral dissertation, University of South Florida, Florida.

Chickering, A.W. & Gamson, Z.F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39* (7), 3-7.

Chizmar, J.F. & Walbert, M.S. (1999). Web-based learning environments guided by principles of good teaching practice. *The Journal of Economic Education*, *30* (3), 248-259.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R. & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9-13.

Dabbagh, N. (2003). Scaffolding: An important teacher competency in online learning. *TechTrends: For leaders in education and training, 47*(3), 39-49.

Knowles, M. S. (1975). *Self-directed learning*. New York: Association Press.

Kouzes, J.M. & Posner, B.Z. (1995). *The leadership challenge: How to keep getting extraordinary things done in organizations*. San Francisco, CA: Jossey-Bass.

Lambert, L. (2003). *Leadership capacity for lasting school improvement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Long, H.B. (2000). Understanding self-direction in learning. In H. B. Long and Associates (Eds.), *Practice & Theory in Self-Directed Learning*, 143-150. Schaumburg, IL: Motorola University Press.

Shavelson, R.J., Phillips, D.C., Towne, L. & Feuer, M.J. (2003). On the science of education design studies. *Educational Researcher, 32*(1), 25-28.

Skill, T.D. & Young, B.A. (Winter, 2002). Embracing the hybrid model: Working at the intersections of virtual and physical learning spaces. *New Directions for teaching and learning,* (92), 23-32.

Sloane, F.C. & Gorard, S. (2003). Exploring modeling aspects of design experiments. *Educational Researcher, 32*(1), 29-31.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press. (Original work published in 1934).

Zaritsky, R., Kelly, A.E., Flowers, W., Rogers, E., & O'Neill, P. (2003). Clinical design sciences: A view from sister design efforts. *Educational Researcher*, 32 (1), 32-34.

Self-Efficacy for Instructional Television Programs of Anadolu University in Turkey

Mujgan Bozkaya
Anadolu University

Abstract

Although large number of available systems, the broadcasted instructional television programs as well as textbooks remain as the main distance education delivery systems in Anadolu University distance education programs. However, the number of studies that investigate either the relationship between prior experience and self-efficacy of the distance students for instructional television programs or the relationship between self-efficacy levels of the students and their satisfaction from those programs is quite few. This presentation involves the results of a study in which the above relationships were examined.

Introduction and Background

Over the last decade the range of media available for distance teaching and learning has increased considerably. However, printed materials and broadcast television programs have remained as the main delivery systems in many distance education applications due to their capability of providing instruction to large numbers of students.

Willis (1993) has indicated that instructional television is an effective distance education delivery system that can be integrated into the curriculum at three basic levels: (1) Single lesson - programs address one specific topic or concept, providing a lesson introduction, overview, or summary; (2) selected unit - a series of programs providing the content foundation for a learning unit in the course curriculum; (3) full course - programs from one or more instructional television series may be integrated into a full semester course typically in conjunction with instructional print materials.

As McIsaac and Gunawardena (1996) have stated, advances in both broadcasting and video formats as well as some other factors have increased the number of distance education providers in academic or corporate settings during 1980s all over the world. Anadolu University is one of those providers has started to offer distance higher education programs to Turkish students by using textbooks and broadcast television programs.

Instructional television has been continuing to be one of the two main delivery systems in Anadolu University's distance programs since the early days of the programs. It has been integrated into the curriculum at the full course level. In other words, for each course in the programs, there is a series of television programs that are used in conjunction with textbooks. All these programs have been produced in-house -Anadolu University has professional television program production facilities- and aired by the Turkish Radio and Television (TRT), a state channel.

Although it has been more than 20 years, there are just a few attempts to investigate the effectiveness of Anadolu University's instructional television programs. Most of those attempts, also, cannot be considered as scientific study. However, they presented several important problems about the programs such as broadcasting time and talking-heads. Obviously there is a need for studies that try to answer questions like "How effective the instructional television programs on distance students' achievement, satisfaction, persistence, attitudes?", "What are the ratings of these programs?", "How the students think about learning from television?", and so forth.

The thoughts or perception of the students about learning form television is the main focus of this study.

Bandura (1986) considers these perceptions as self-efficacy beliefs of learners and defines as "a judgment of one's ability to execute a particular behavior pattern" (Bandura, 1978, p.240). According to this definition self efficacy beliefs form a central role in regulatory process through which an individual's motivation and performance attainments are governed (Wood & Bandura, 1989). On the other hand, self efficacy judgments or beliefs determine how much effort people will spend on a task and how long they will persist with it. But it is not concerned with the strategies one has but with judgments of what one can do with whatever strategies one possesses. Increasing number of researchers (e.g. Joo, Bong & Choi, 2000; Lee, 2002; Multon, Brown & Lent, 1991; Pajeres, 1996; Schunk, 1994) have been investigating the relationship of self-efficacy to learning and academic achievement, but research in the area of academic performance and satisfaction is still developing.

In the literature, there are some evaluation studies about learning from television but no study that looks for the relationship between the levels of self-efficacy for learning from television and satisfaction can be found. This sort of a study might provide interesting data to the providers of distance education and the designers of the distance learning environments such as the Anadolu University.

Furthermore, according to Bandura (1997), the four major sources that contribute to an individual's self-efficacy are enactive experiences (success or failure), vicarious experiences (observing others' success or failure), verbal persuasion (of spouse, friend, etc) and physiological states (anxiety, fear, etc). The relationship between distance learners' enactive experiences for course content or specific delivery systems and their academic achievement or satisfaction has been examined over the years. Some of these studies show that when learners have more experience their self-efficacy increases while in others no relation is found.

In Turkey as well as other countries, no study that seeks the relationship between past experience and self-efficacy for learning from television has been found.

Purpose

This study is mainly seeks to determine if there is relation between self-efficacy and satisfaction levels of Turkish distant students for instructional television programs. The research questions of the study were formulated as:

1. What are the self-efficacy and satisfaction levels of the students for learning from television programs?
2. Do the self-efficacy and satisfaction levels of the students differ according to gender?
3. Do the self-efficacy and satisfaction levels of the students differ according to prior experiences of learning from television?
4. Is there any relationship between distance students' levels of self-efficacy and satisfaction?
5. What do the students think of improving the effectiveness, efficiency and appealing of the instructional television programs of the Anadolu University?

Methodology

The study conducted with the students enrolled the distance education programs of the Anadolu University. Two scales – a self-efficacy for learning from broadcast television programs and a satisfaction – were developed for the study. The details of the subjects, the instruments and the procedure were given below.

Subjects

The subjects of this study are the students who are enrolled in Anadolu University's distance learning programs, live in the city of Eskisehir, and attend the face-to-face evening classes in the University campus. Although 250 students accepted to participate the study and filled out the self-efficacy instrument at the beginning, only 72 of these students were participated the second part (satisfaction) of the study. Of these students 47 were female and 27 were male.

Instruments

Two instruments have been developed for this study. In order to find out the self-efficacy levels of the students for learning from television, a 5-point instrument, which is actually adaptation of several self-efficacy scales for varying delivery systems such as the Self-Efficacy for Online Technologies Scale of Miltiadou (2000) and the Computer Self-Efficacy Scale of Cassidy and Eachus (2001), has been developed. The researcher developed a list of constructs to be measured before composing the items. These constructs for learning from television were effectiveness, efficiency and appealing of the programs, and learning from television. A review committee reviewed the final draft of measures and revised some of the items. At the end, the Self-Efficacy for Learning from Television instrument consisted of 24 items. Of these 24 items 7 were related to effectiveness, 4 were to the appealing, 5 were to efficiency of the television programs, and other 8 items were associated with learning from television.

The 3.41 mean score identified as the expected level of self-efficacy with the item while other responses enables students to show higher or lower levels of self-efficacy. The 3.41 mean average was determined after identifying the critical level: 4 intervals/5 categories = 0.8. As a result of this the levels of self-efficacy were determined as followings:

1	= 1+ 0.8	= 1.8
2	= 1.8 + 0.8	= 2.6
3	= 2.6 + 0.8	= 3.4
4	= 3.4 + 0.8	= 4.2
5	= 4.2 + 0.8	= 5

To determine the satisfaction levels of the students, a satisfaction instrument has been developed. In the process of developing this instrument, Gunawardena and Zittle's Satisfaction Scale (1997) and Biner's Satisfaction Questionnaire (1993) have been used. The instrument includes 10 items and required students to choice from a 5-point scale. For this instrument, the 3.41 mean score is also considered as the critical level of satisfaction.

In addition to these instruments, an open ended question asked the students to gather data about the students' thoughts and recommendations on improving the effectiveness, efficiency and appealing of the television programs.

Procedure

The self-efficacy scale was administered the 250 students in January 2003 just before the instructional television programs started to be broadcasted. Unfortunately almost all of these students were first year students (freshmen) and no third year student (junior) accepted to join the study. So the third research question of the study has been discarded.

It was planned to administer the satisfaction instrument to the same students at the end of the semester, July 2003. Unfortunately, only 72 students attended the second part of the study. So it was conducted with those students. During the second part of the study, the students were also asked to answer the open ended question about the improvement of the television programs.

All the responses (in paper format) of the subjects were then coded according to the developed coding system. SPSS V10.0 was used to conduct all the statistical analysis.
All the procedures were adopted to establish the construct validity for the instruments. The reliability of educational television programs self-efficacy scale was found as 0.92 and satisfaction scale was found as 0.90.

Findings

As it has been mention above the 3.41 mean average was taken as the critical level of self-efficacy and satisfaction in interpreting the results.
Self-Efficacy and Satisfaction Levels. Average mean of students' responses to items concerning their perceived self-efficacy levels for learning form television was lower then critical level (M_{se}=2,98 > M_{cl}=3,41). For all constructs of the instrument (effectiveness, efficiency, appealing, and learning) the average means of students' responses were lower then critical level (Table 1). This result can be interpreted as that the distance students did not believe in that they could learn from television programs at the beginning of the semester.

Table 1. *Mean Scores of Students' Responses*

Instruments / Constructs	Mean Scores	
Self-Efficacy	2,98	
Effectiveness		2,86
Efficiency		2,78
Appealing		3,15
Learning		3,15
Satisfaction	2,79	

On the other hand, average mean scores of students' responses to items related to the learning from television and appealing of the programs were same (M_{lr}=M_{ap}=3,15) and higher then other average scores of the responses to items concerning effectiveness (M_{efv}=2,86) and efficiency (M_{efc}=2,78). Also this score is quite close the critical level (M_{cl}=3,41). By looking at these scores, it can be told that the students find television as a useful medium for learning, believe it might help them pass the courses, and like to learn from television.

Additionally, the students scored higher on the item #13 "I believe that a picture worth of billion words" (M_{13}=3,35) and item # 24 "I think television is quite a proper medium for learning" (M_{24}=3,31). All these items were considered under the learning from television construct.

Moreover, one of the items that got the lowest score in the instrument was item #4 "I believe that trying to learn a topic from television is just a waste of time" (M_4=2,49). This item coded reverse during the analysis and classified under the efficiency construct. Another item that also got the lowest one of the lowest scores was item #3 "I feel that I can do better at the exam when I learn the topic from television" (M_3=2,64). This item was considered as one of the effectiveness items.

Average mean of students' responses to items in the satisfaction instrument was also lower than the critical satisfaction level (M_s=2,79> M_{cl}=3.41). This result shown that the students did not satisfied learning from television programs (Table 1).

Gender. In order to see the difference between female and male students' responses to the instruments, *t*-test analyses were used. Table 2 illustrated the mean scores of female and male students in both tests and the comparison of these scores.

Table 2. *Mean Scores of Female and Male Students in both Instruments and Results of the t-test Analyses*

Instruments Constructs	**Sex**	**Mean**	**N**	**t**
Self-Efficacy	Female	2.95	47	.74
	Male	3.04	27	
Effectiveness	Female	2.76	47	.44
	Male	2.83	27	
Efficiency	Female	2.84	47	.47
	Male	2.90	27	
Appealing	Female	3.10	47	.68
	Male	3.24	27	
Learning	Female	3.10	47	.68
	Male	3.24	27	
Satisfaction	Female	2.85	47	.75
	Male	2.69	27	

*p< .05 **p<.01

The comparison has shown that the mean scores of female students were lower than that of male student in self-efficacy instrument including all the constructs. However no significant difference related to gender has been noticed. On the other hand, in the satisfaction instrument female students scored higher than male students. But, same as self-efficacy instrument, this difference was not statistically significant, too. In the light of these analyses, it can be told that gender does not make any difference in the students' self-efficacy and satisfaction levels.

Prior Experience. Unfortunately no third year students accepted to join the study. So, no data has been gathered to see whether prior experience makes a difference in the students' self-efficacy and satisfaction levels for learning from television.

Relationship between Self-Efficacy and Satisfaction. A Pearson correlation analysis was conducted to examine the relationships between the students' levels of self efficacy and satisfaction for learning from television. The result of this analysis is given in Table 3. According to this analysis there was a significant positive relation between the students' self-efficacy and satisfaction levels (r=0,372, N=74, p<0.000, two-tailed). This result can be interpreted as that learning the students' self-efficacy levels for learning from television can be used as one of the indicators of students' satisfaction for learning from television.

Table 3. *Correlation Coefficients between Students' Self-Efficacy and Satisfaction Levels*

		Self-Efficacy	**Satisfaction**
Self-Efficacy	Pearson Correlation	1,000	,372**
	Sig.(2-tailed)		,001
	N	74	74
Satisfaction	Pearson Correlation	,372**	1,000
	Sig. (2-tailed)	,001	
	N	74	74

** Correlation is significant at the 0.01 level (two-tailed)

Improvement of the Television Programs. During the administration of the satisfaction instrument, the students were also asked to answer an open-ended question intended to gather data about their thoughts and recommendations on improving the effectiveness, efficiency and appealing of the television programs. The results were summarized in Table 4.

Table 4. *The Students' Thoughts on Improving the Television Programs*

Thoughts & Recommendations	Frequency
More drill and practice should be employed in the programs	15
The length of each program should be more than 20 minutes	13
The programs should cover more topics	12
Each program should be repeated several times	12
The broadcasting times should be more convenient for the students	12
More visuals should be used in the programs and they should be organized more attractively	10
Visuals (tables, figures, questions, solutions) should stay longer on the screen	10
The language should be more easily understandable	7

The most frequently expressed thought was related to employment of more drill and practice activities into the programs (f=15). This thought can easily be related to the situation in Turkish education system which was elaborated in the discussions. Beside this thought, the ones related to the visuals and the language used in the programs can be considered interesting.

Discussions and Conclusions

According to the self-efficacy and satisfaction levels of the students, it can be drawn that the distance education students of the Anadolu University find the television programs unsuccessful in terms of helping them learn. They found television programs as ineffective and inefficient although they mentioned that television was a proper medium for education. This result can be related to Salomon's claim about television. Salomon (1984) indicated that television is considered as an "easy" medium. The students responses to the item #5 of the self-efficacy instrument, "I think learning a topic from television require lesser effort" also support the Salomon's idea.

In addition, no significant difference has been found between female and male students' self-efficacy and satisfaction levels. This result is consistent with a previous study (Bozkaya, 2002) conducted to learn self-efficacy levels of students in the Anadolu University. In terms of gender, it was interesting to observe that male students had a higher score on self-efficacy than girls but got a lower score on satisfaction.

The study has also revealed that there is a positive relationship between self-efficacy and satisfaction levels of the students. This result supports the literature. In other words, the higher self-efficacy level of an individual related to accomplish a task, the more s/he will be satisfied or vise versa. In this study, the students' self-efficacy level has been identified as low before the television programs aired. Later, at the end of the broadcastings, their satisfaction level was also observed low.

The participant students' answers to the open ended question have provided an idea why the satisfaction level was low. The programs usually designed in way so that they cover the areas that could not included into textbooks. Varieties of methods including role playing, demonstration, field trips, and lecture are employed in those programs to help students acquire required knowledge, skills and attitudes. However, students would like to watch more exam-oriented programs. In other words, they prefer programs that aim to prepare them to the tests. This preference can easily be related to one of the important issues of the Turkish education system. Having a higher education diploma is very important in terms of living a decent life in a country like Turkey where there is a big gap between poor and rich. Thus, every year more than 1,5 million young people take the university entrance exam. Only 1/3 of those people can get in a higher education institution due to lack of demand (not enough number of higher education institutions). Since the university entrance exam is very important, students start preparing themselves to the exam from early grades. Even schools administrations, teachers, and parents force students to prepare themselves for the exam. So the main goal of the system is now doing better than others in the university entrance exam. In other words, the goal is not learning but passing the exams. Same situation exists in distance education programs too. Most of the students' main goal is passing the exams and having the diploma but not learning new skills and knowledge. Most probably that is why students prefer television programs intend test-preparation rather than instruction.

As a conclusion, this study has shown that identifying students' self-efficacy levels for learning from television might help to predict their satisfaction from instruction television programs in distance education settings. In order to increase the satisfaction, the designers might focus on students' needs and wishes. However, this does not mean that the main goal should be test-preparation. The main goal must always be helping students learn but a balance must be attained between what should be taught and what students want.

Additionally, it might provide interesting data if a similar study can be conducted with a large number of subjects including enough number of third and fourth year distance education students. So that relationship between prior experience and self-efficacy levels of the students for learning from television would be examined.

References

Bandura, A. (1978). Reflections on self-efficacy. *Advanced in Behavioral Research and Therapy,1*(4), 237-269.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H.Freeman and Company.

Bozkaya, M. (November 2002). The Affects of an Online Course on Turkish Undergraduate Students' Perceived Computer Self-Efficacy and General Self-Efficacy. Paper presented in *AECT 2002 International Conference*, Dallas.

Biner, P.M.(1993). The development of an instrument to measure student attitudes toward televised courses. *The American Journal of Distance Education, 7*(1), 62-73.

Cassidy, S. & Eachus, P. (2001). *Developing the computer self-efficacy (CSE) scale: Investigating the relationship between CSE, gender and experience with computers*. Retrieved February 2002 at http://www.salford.ac.uk/healthSci/selfeff/selfeff.htm

Gunarwardena, C. N. and F. J. Zittle. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American Journal of Distance Education, 11,* 8-26.

Joo, Y., Bong, M. & Choi, H. (2000). Self efficacy for self-regulated learning, academic self-efficacy and Internet self-efficacy in web-based instruction. *Educational Technology Research and Development, 48*(2), 5-17.

Lee, I. S. (2002). Gender differences in self-regulated on-line learning strategies with Korea's university context. *Educational Technology Research & Development, 50,* 101-111.

McIsaac, M. S., & Gunawardena, C. (1996). Distance education. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (p. 148). New York: Simon & Schuster Macmillan.

Miltiadou, M.,& Yu, C.H. (2000). Validation of the online technologies self-efficacy scale (OTSES). *International Journal of Educational Telecommunications*.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30-38.

Pajeres, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543-578.

Salomon, G.(1984). Television is " Easy" and Print is "Tough": The Differantial Investment of Mental Effort in Learning as a Function of Perceptions and Attributions. *Journal of Educational Psychology,76*(4), 647-658.

Schunk, D.H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In D.H. Schunk & B.J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 75-99). Hillsdale, NJ: Erlbaum.

Willis, B. (1993). *Distance education: A practical guide*. Englewood Cliffs, NJ: Educational Technology Publications

Wood, R. & Bandura, A. (1989). Social Cognitive Theory of Organizational Management. *Academy of management Review,14*(3), 361-384.

Return On Investment in Nature Conservation—Evaluating Leadership Development in the World Wide Fund for Nature

Jeroen Breman
WWF College for Conservation Leadership
The Netherlands

Diederick Stoel
Profitwise
The Netherlands

Abstract

New challenges in conserving the world's biodiversity created a need for a new type of leadership in the World Wide Fund for Nature. The WWF College for Conservation Leadership was founded to address this need. An 18-months programme combining two face-to-face workshop with online distance learning was established. This paper describes an evaluation study of the programme aimed at measuring the impact on organizational results.

Background

The World Wide Fund for Nature (WWF) is one of the world's most active and professional conservation organizations. Since 1961, WWF has been involved in programmes around the world to try and restore the balance between human activity and nature. Over those 40 years, WWF and other environmental groups can claim some outstanding successes. But, the bottom line is that while WWF may be winning some battles, they are certainly losing the war. The Living Planet Index (2002), which tracks macro-level global biodiversity, shows that the world's species and habitats have declined 35% over the past 30 years. The world will be biologically poorer tomorrow than it is today. The blunt truth is that we must do more.

Conservationists face increasingly complex and large-scale challenges. They need to excel in technical areas, and also need to be able to engage and motivate partners: civil society, local and national governments, business and industry, the media and local communities. WWF needs to develop and motivate their teams and networks. In short, conservationists need a complex set of skills to carry out our jobs effectively. Above all WWF needs leadership. As one College participant put it, 'This is not just a biologist's job anymore'.

The WWF College for Conservation Leadership was conceived in 1998 to help WWF and its partners meet these challenges and to develop the leadership skills needed to start winning the war.

Programme Design

The WWF College for Conservation Leadership aims to enable individuals to recognize, challenge, and develop leadership talent in themselves and others in order to improve the impact of nature conservation. The College strives to do this by delivering individualized online learning programmes, and creating learning environments to share challenges, knowledge, and approaches.

The College is an interactive learning network. It combines new technology with traditional face-to-face training and offers a unique learning environment. Participants learn from each other, and analyze and improve their own performance to become better leaders and thereby achieve greater conservation results. The College promotes the ideas of self-directed learning, peer learning, learning on the job and continual improvement.

The Conservation Leadership Programme is an 18-month programme with modules around four themes that have been identified as important for leaders in nature conservation: *Leadership and Management*, *Strategy and Planning*, *Campaigning and Advocacy*, and *Communication and Networking*.

The programme is designed and developed in such a way that it can easily be part of the workday, with a time investment of an average of two to four hours a week, and modules that can directly influence the way that participants do their work. The College brings people together from around the WWF Network to share experiences, discuss new trends, and help develop the competencies that make up a leader in conservation. The College is not a just a place were participants take. They already have a lot of knowledge and skills which they bring to the College. The College is THE place to learn, share, and lead!

Each year, through a competitive selection process that involves both the applicants and their supervisors, two groups of 25 participants, balanced in terms of geographic regions, gender and functional specialties, start the Conservation Leadership Programme. The programme begins with a one week workshop which sets the stage for the rest of the 18 months. Through role plays, exercises and discussions, this interactive workshop introduces the themes of leadership, communications, and information communications technology. Participants evaluate themselves on the key competencies within each of the four themes, and then in consultation with a facilitator, plan a customized study programme based on their needs. This week also provides an opportunity to get all participants familiar with the College' online learning environment, the College Campus.

A primary task of the facilitators (both in the workshops as well as online), is to ensure a two-way flow of learning. Rather than having an expert stand in front of the group and lecture on the 'right' way to do things, participants are encouraged to apply what they know and share ideas with other participants to learn from each others' experiences, add some fresh perspectives to their own point of view, and thereby increase the skills in their toolkit. This has the added benefit of building a greater capacity for learning and problem-solving by using their networks in the future. Participants are exposed to new theories and models that they are then asked to apply in their daily work situations.

Following the opening workshop the programme continues with nine months of online distance learning using the College website. Participants select and take part in a series of six-week modules, requiring 2-4 hours of study per week. These modules are facilitated by a content expert, who reviews online assignments, answers questions and moderates discussions. Participants must satisfactorily complete all assignments to finish a module. There are two core modules, Leadership Skills and Effective Communications, that all participants have to complete. Next to the core modules, participants need to complete at least five elective modules that can be selected from a list of sixteen modules spread over the different themes.

The group comes together again mid-way through the programme, for another week-long workshop. This is an important time to review progress, revise the study programme as needed, and address any problem areas. Also during this interactive workshop participants work on leadership, teambuilding, strategic partnerships and campaigning through role plays, exercises and a group project based on a real-life case study from the region of the workshop location.

This second workshop is again followed by a further nine months of online learning. At the end of the 18 months, successful candidates receive a Certificate of Conservation Leadership.

The College started with a pilot phase of four groups. Even before Group 4 finished, WWF's senior management decided to continue the programme. Since then, five more groups have started their 18-month programme. Of the first four groups with a total of 70 participants, 30 successfully completed the programme. Of the participants that started and finishes since the pilot phase, 75% successfully completed the programme. A handful others had valid reasons to ask for extension of their programme and may graduate in the next couple of months. Reasons for allowing extensions are amongst others: health problems, family crises, and module scheduling conflicts.

Purpose of the Evaluation

In an ideal world, evaluation is planned before or during the design of a programme. As in many cases, this was not done in the design of the Conservation Leadership Programme.

Until now, reporting of successes in the College was very activity based. For instance locations of workshops, numbers of participants starting their programmes and how many complete the programme is reported. Some post-module and post-workshop satisfaction ratings are available, but little is known about the quality of the programme—or more important—of the participants' learning. And are new knowledge, skills, and attitudes actually used? For the College management, the main questions are: a. are the right things being done? and b. are things done right?

At the same time, the global trend of accountability of all kinds of processes is not surpassing non-profit organizations such as WWF. Senior management of the organization is starting to wonder what the impact is of sending participants to and funding a training programme. Can the impact of a leadership training on nature conservation be measured?

This evaluation study is a first attempt in WWF's Conservation Leadership Programme. The main questions in this evaluation study are: a. are participants learning the offered knowledge and skills? b. are participants applying them in their work? c. what is the impact on WWF 's critical organizational measures? and d. can the impact be translated in monetary return?

Method

In the following paragraphs the method of the evaluation study is describe. First, the theoretical background is explained. Next, the translation is made to the current study: subjects and instruments are described. Finally, the mechanism for isolating the effects of the programme on the result is clarified.

Measuring ROI of human resources, training, and performance improvement programmes is becoming common in businesses around the world. Increased emphasis on understanding inefficiencies in other sectors such as the health-care, non-profit, and public sector, but also building accountability into their processes causes these sector to adopt the ROI process (Phillips, 2001).

This study used a five-level evaluation process, based on Kirkpatrick's four levels of evaluation (Kirkpatrick, 1994), and Jack Phillips' return-on-investment (ROI) process (Phillips, 1997). This process enables the measurement of reaction and learning (Level 1 and 2) at the end of workshops and modules. It also enables the measurement of job application of knowledge and skills (Level 3) and organizational results affected by the

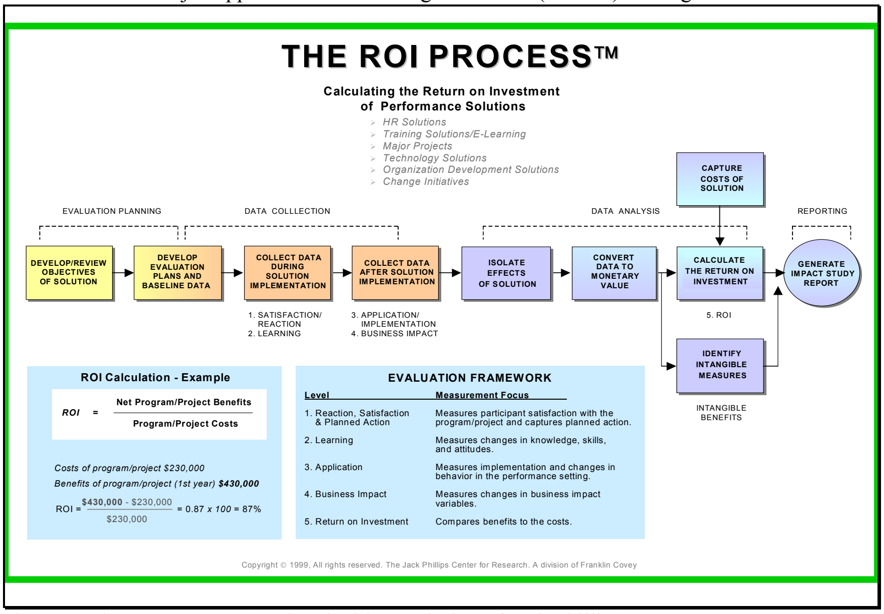


Figure 1. The ROI process (reprinted with permission of Jack Phillips)

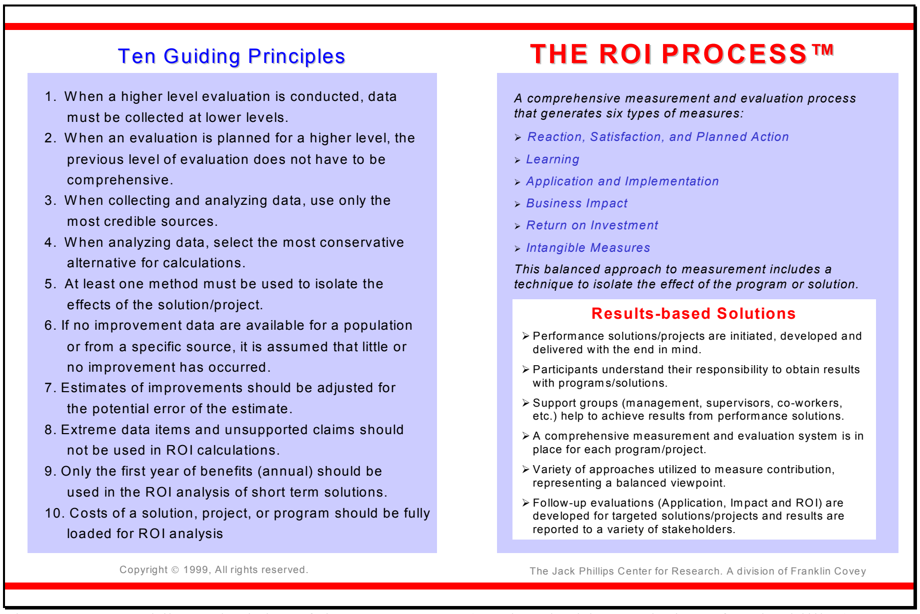


Figure 2. Ten guiding principles of the ROI process (reprinted with permission of Jack Phillips)

programme (Level 4). In addition the process offers a tool to calculate the programme's ROI (Level 5). Phillips (1997) emphasizes that the higher levels can not be seen and measured without paying attention to the lower levels.

Evaluation according to Phillips (1997) is much like putting together a puzzle with five pieces: a. The evaluation framework, which defines the levels at which programmes are evaluated as described above. b. The

ROI model. Figure 1 shows the ROI methodology model, depicting systematic steps to ensure consistent application of the methodology. c. Operating standards and guidelines (see Figure 2) that build credibility in the process and support a conservative approach to ROI evaluation. d. Practice provides support for the final piece: e. Implementation, bringing together the pieces of the puzzle. As mentioned before, the evaluation of the Conservation Leadership Programme was not planned when designing and developing the programme. Thus no assessment of knowledge and skills before participation was collected. Still the ROI process used in this study provides useful tools to measure learning, application and impact after the programme. And management of the College decided that this first effort is preferred over no effort to evaluate.

Subjects

All 44 participants from Group 4 and 5 of the Conservation Leadership Programme were included in the study. The participants, 29 male and 15 female, are spread over 27 countries in Asia, Africa, Central and South America, and Europe. They have a variety of functions in the organization, such as communications, finances, human resource management, and different functions in nature conservation units and projects. Group 4 graduated one year before receiving the questionnaire. For Group 5 this was three months. The groups were the most recent graduates of the programme.

Instruments

Reaction data are continuously collected after each workshop and each online module. In the case of a workshop on paper, after an online module with an online form. Some focus on learning of new skills is included in these forms, but questions about Level 2 were also included in the questionnaire that is described next.

A questionnaire was developed to measure results at Levels 2 through 5. On a five-point scale, participants were asked to identify their degree of success in achieving the learning objectives for the different competencies underlying the programme (Level 2) and list the top five most frequently used competencies. For Level 3, application of skills, participants were asked to indicate on a five-point scale the perceived change in a list of skills that are used in the College to characterize the different competencies. This level also included a top-five of skills that impacted their work most. For impact (Level 4) a list of measures was developed in cooperation with the Steering Group of the College. This group—in the College called 'Theme Owners' is among others responsible for guarding the quality of the programme in the different themes of the College programme, and consists of senior management members from across the WWF-Network. Measures were combined in 8 categories: financial income, financial cost savings without loss of conservation impact, Communication and media, Campaigning, Networking and partnerships, Policy change, Conservation, Innovation and learning. Within these categories, indicators focus on hard data such as output, quality, costs, and time.

For ROI (Level 5), two types of data are necessary, programme costs and programme benefits. Costs for the participants include: salary costs for the time spent on the programme including benefits, tuition, and travel, accommodation, and meal costs for two workshops. For benefits, participant were asked to reflect on the indicators of the impact question and select three or more accomplishments that can be converted into monetary value.

After two months and two reminders, a short survey was sent to non-respondents. In this survey they were asked to indicate what the main reason was why did not manage to return the questionnaire.

Isolation method

The most critical step in the evaluation process is one that is often overlooked: isolation of the effects of the programme (Phillips, 1997; Phillips, 2001). What part of the impact can be attributed to the intervention, in this case the Conservation Leadership Programme, and what other influences may underlie the results? Phillips (1997) and Phillips (2001) describe several techniques that have been used to address the issue. In the present study, participant estimates of programme impact have been used.

For each benefit listed in the questionnaire, participants were asked to estimate the percentage of the improvement that was actually influenced by the application of skills learned in the Conservation Leadership Programme. Also, participants were asked with what level of confidence they made that estimate. For the calculation of ROI the most conservative estimate of the monetary value was used.

For example, if a participant estimated a monetary benefit of $4,000 and 60% of the improvement is due to the Conservation Leadership Programme. The participant is 80% confident about that estimation. Then the most conservative estimation of the benefit will be 60% x 80% x $4,000 = $ 1.920.

Results

Of the 44 questionnaires, 15 were returned. Only nine of the participants answered all questions. The other six answered the questions up until the questions on monetary values. Participants used comments like "hard to estimate", "there is no reason to calculate this", and "most of it is qualitative but still important" as arguments to stop their answers at Level 4. The difference in time between the end of the programme of the two groups and filling out the questionnaires did not show a significant difference in the results. Therefore the results are combined.

In general participants react positive to this programme. One skill that is taught in the Conservation Leadership Programme is to be able to provide colleagues and staff with constructive feedback. Participants learn this skill during the first workshop and will use it in the evaluation forms. In Table 1, the average ratings of the main themes of the workshops are given for both groups. It is not possible to list all evaluation results of all modules.

Table 1. *Selected reaction data for College workshop main themes*

Kick-off Workshop	Group 4	Group 5
Leadership	3.72	4.57
Communication	4.47	4.75
Vision	4.11	4.63
ICT	4.83	4.21
Follow-up Workshop		
Strategic partnerships	4.32	4.20
Networking	3.90	4.24
Campaigning	4.55	4.50
Teambuilding	4.00	4.45

Participants report valuable learning across the themes of the programme. The ten competencies with the highest degree of success to reach the learning objectives are listed in Table 2. The lowest average score on any of the items in this questions was 3.7. In Table 3 the competencies applied most on the job are listed. Motivating and directing individuals and themes, a competency within the Leadership and Management Theme of the programme was mentioned by eleven of the fifteen respondents. Vision and Strategic partnership each by ten participants.

Table 2. *Top 10 Level 2 data: Degree of success in reaching learning objectives*

Competency	Score
Capitalizing on knowledge sharing	4.4
Developing and communicating a vision	4.3
Using communication to procure results effectively	4.1
Motivating and directing individuals and teams	4.0
Mobilizing the Network and other NGOs	4.0
Recognizing and using media opportunities	4.0
Managing a project through all its cycles	3.9
Building and managing strategic partnerships	3.9
Creating awareness and educating others	3.9
Using communication technology effectively	3.9

Table 3. *Top 3 competencies applied most on the job.*

Motivating and directing individuals and teams
Developing and communicating a vision
Building and manage strategic partnerships

When looking at the skills applied on the job (see Table 4), the skills learned in modules in the Leadership and Management Theme have improved most. The six most improved skills are found in this theme, before skills in Strategy and Planning, and Campaigning and Advocacy. In Table 5 the four skills that had the biggest impact on their job performance are listed. Here also leadership skills were reported most, followed by campaigning and advocacy skills.

Table 4. *Top 10 Level 3 data: Level of improvement of applied skills*

Theme	Skills	Score
L&M	Establishing clear directions and objectives	4.1
L&M	Inspiring and motivating	3.9
L&M	Finding common ground and getting cooperation	3.9
L&M	Formulating and communicating a vision	3.8
L&M	Empowering others	3.8
L&M	Sharing wins and successes	3.8
S&P	Maintaining a positive relationship with potential partners	3.7
C&A	Presenting case clearly	3.7
L&M	Analyzing strategically	3.7
L&M	Communicating clearly	3.7

L&M: Leadership & Management; S&P: Strategy and Planning; C&A: Campaigning and Advocacy.

Table 5. *Top 4 skills having most impact on the job.*

L&M	Analyzing strategically
L&M	Formulating and communicating a vision
C&A	Using relevant communication based on target audience
C&A	Developing and implementing appropriate campaigns

Table 6. *Top Level 4 data: Degree of influence of training on organizational measures*

Category	Measures	Score
I&L	Knowledge and experience shared with team members	4.2
C&M	Quality of message – style, format, packaging	3.8
C&M	Reach – number of target audience exposed	3.3
I&L	Time spent on training and development	3.2
C&M	Continuity – distribution of messages over the year	3.1
N&P	Dialogue with potential partners established	3.1
I&L	Knowledge and experience shared on the web	3.0
N&P	Strategic partnerships established	2.9
CON	Magnification achieved	2.9
CON	Conservation Capacity	2.9

I&L: Innovation & Learning; C&M: Communication & Media; N&P: Networking & Partnerships; Con: Conservation

Table 7. *ROI Calculation*

	Group	Benefits	Costs	CBR	ROI
1	4	-	15,676.32	-	-
2	4	-	16,898.40	-	-
3	5	-	13,826.00	-	-
4	4	15,200	6,312.80	2.41	141%
5	5	38,321	14,108.36	2.72	172%
6	5	20,000	15,330.00	1.30	30%
7	4	-	14,710.40	-	-
8	5	-	11,290.00	-	-
9	5	67,500	24,335.91	2.77	177%
10	5	56,290	15,663.70	3.59	259%
11	5	-	13,882.00	-	-
12	5	4,955	12,300.00	0.40	-60%
13	5	14,155	6,434.60	2.20	120%
14	4	34,347	6,940.00	4.95	395%
15	4	8,000	7,705.00	1.04	4%
Total		258,768	195,413.00	1.32	32%

CBR: cost-benefit ratio.

For Level 4 data, the participants were asked to indicate to what extent they thought that application of knowledge and skills had a positive influence on organizational measures. Measures in eight categories were listed. Table 6 shows the measure with the highest degree of influence. Results show that items in the Innovation and Learning, and the Communication and Media categories score best followed by indicators in Networking & Partnerships, and Conservation.

In Table 7 the calculations for the ROI are given. For each of the fifteen participants the costs and benefits that they reported are listed. Next to the ROI the cost-benefit ratio (CBR) is given. CBR is a different way of calculating return. The CBR is calculated by dividing costs by benefits. ROI is costs minus benefits divided by benefits. For the six participants who did not report a benefit, a ROI could not be calculated. By default their ROI is -100%.

The figures are a representation for the ROI of the individual participants. The figures in Table 7 do not include development costs and other costs the College has. For the individuals working in their individual offices and projects these are their real costs. When College overhead costs are not included, the ROI for this group of participants is 32%.

For WWF as an organisation the programme costs should be included. The tuition paid by the participants does not cover all costs. The organisation invests in the developments of its workforce. When these costs are calculated per participant, the ROI for the fifteen respondents would be -29%, or for the whole group of forty four participants -76%.

Table 8. *Non-respondents survey*

Reason for not responding	Frequency
no time	18
other engagements	7
other priorities	4
the questions are too difficult to answer	5
the questions are irrelevant	0
I could not open the document	1
I didn't receive the questionnaire	3

Twenty three out of twenty nine non-respondents, reacted on the survey sent two months after the questionnaire was distributed. More than one answer was possible. The main reason for not responding was time. Only a few participants thought the questionnaire was too difficult.

Discussion and Conclusion

The main questions in this study were if participants in the College are learning and applying the necessary knowledge and skills from the programme, what the impact of this learning on the organization is, and if ROI can be calculated for these impacts.

Participants reported learning in all competencies of the Conservation Leadership Programme. On a five-point scale, participant rate their own learning between 3.7 and 4.4 on 13 different competencies. When comparing the competencies that participants list as being used most on the job (see Table 3) and the level of reaching the learning objectives of these competencies (see Table 2), it becomes clear that the most used competencies not always score the highest in the degree of success of learning. Developing a vision for conservation and reaching conservation targets through strategic partnership are currently emphasized as two of the main strategies for achieving WWF's goals. The difference in scores in Table 2 are not very big. Further analysis is necessary to determine if the content of the programme should be adjusted to increase learning in these key areas.

In the description of the background of the College, it was mentioned that "above all, WWF needs leaders." The results of this study suggest that participants have improved their leadership and management skills in their daily work. Eight skills in the Top 10 Level 3 data in Table 4 are skills learned in modules from the Leadership and Management theme. The Conservation Leadership Programme appears to achieve its mission in developing leadership talent.

The biggest impact of the College programme on achieving WWF's key organizational measures are knowledge sharing and quality of communication (see Table 6). In the Top 10 list, many learning and communication indicators score high. Conservation measures appear in 9-th and 10-th position in this list. The two categories listing financial measures (income, and cost savings) do not appear at all in the top, just like policy change, and campaigning.

The fact that the participants were part of a knowledge sharing experience influenced this score. And in the College programme, communication is the topic of one of the mandatory online modules. This could explain the high score in these measures. Financial benefits and savings may not be so obvious for the participants. Applying for funds from either internal or external sources takes time. Also budgets are prepared once a year and there are limited possibilities for revision. This may cause that impact is hard to measure on the short term, but could still have an effect after a year. Still the College should further investigate if the programme can be improved to support the impact on the less influenced measures.

The question if ROI can be calculated should be answered with yes. However it is not for everyone in the group of respondent an easy task. Participants who reported financial benefits from the programme can explain well what the basis is for their estimation. For example: new products developed, project proposals approved, new strategic partnerships, time savings, cancellation of projects, and effects of new training programmes implemented. When asked what percentage of these benefits are due to participating in the College programme, participants' estimates vary from 30% to 100% with an average of 72%. They are very confident about their estimations. The level of confidence varies from 70 to 100% with an average of 88%.

The nine participants reporting a return in financial terms report benefits that are sufficient for the whole group of 15 respondents. For every dollar that the fifteen spend on participating in the College 1 dollar and 32 cents was returned.

The question remains what distinguishes the nine participants who did report a return in financial terms from those who failed to do so. With current economic uncertainties and an increasing emphasis on accountability, the ability to clearly indicate the impact of certain processes on achieving organizational targets could show to be an essential competency that every Conservation Leader should have. More research will be needed to identify the differences between the two groups.

References

Kirkpatrick, D. L. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler Publishers.

Phillips, J.J. (1997). *Return on investment in training and performance improvement programs*. Boston, MA: Butterworth-Heinemann.

Phillips, P.P. (Ed.). (2001). Measuring ROI in the Public Sector. In, J.J. Phillips (Ed.), In action series, Volume 3. Alexandria, VA: American Society for Training and Development.

World Wide Fund for Nature. (2002). *Living Planet Report 2002.* Online publication: http://www.panda.org/news_facts/publications/general/livingplanet/index.cfm

Reaching the Teachers: The Key to Technology Adoption in the Schools

Janet A. Buckenmeyer

Abstract

While research suggests that conditions exist which do influence teacher use of computer technology, it is not known which conditions or combination of conditions are more likely to influence or be related to the secondary teacher's decision to adopt selected educational technology in the curriculum. This purpose of this study was to determine the strength of the relationships, if any, between: (1) physical resources and technology adoption by teachers in schools, (2) teacher attitude and technology adoption by teachers in schools, (3) staff development and technology adoption by teachers in schools, and (4) demographics and technology adoption by teachers in schools.

This study derived its theoretical constructs partly from theory on change process, including Rogers' (1995) Diffusions of Innovations and the Concerns-Based Adoption Model. A survey instrument was drawn and modified from various sources in order to assess the strength of the potential relationships listed above. Secondary teachers (N = 144) from a local suburban school district participated in the study. Correlations with technology adoption were found to be significant with: (1) staff development, and (2) physical resources.

Five implications emerged from this study: (1) provide staff development opportunities, (2) provide time, (3) provide prompt technical assistance, (4) provide incentives, and (5) promote positive attitudes toward technology.

Introduction

Various educational technologies have been introduced within the educational environment over the last half-century. From phonographs and audiocassettes to instructional television, all educational technologies have intended to change the educational scene. However, the most pervasive and far-reaching of these influences has been the introduction of the computer and related technology into the classroom (Office of Technology Assessment, 1995). While not offering a panacea to the ills that plague our nation's schools (OTA, 1995), computer technology does provide our teachers and students with an overabundance of opportunities that were not only unavailable in previous generations, but were not possible with the aforementioned technologies. Such teaching advantages with computers include simulations of otherwise costly, time consuming, or dangerous situations, instant global communications and immediate access to limitless amounts of information. In spite of this potential, school administrators and teachers in the United States have been resistant in adopting this technology. Studies continue to show that beginning teachers still do not have the skills needed to successfully integrate technology into the curriculum (Firek, 2003). A national survey by Becker (1999) supports the finding that a majority of teachers, as many as seventy percent, are not using the technologies available to them. Further, in *Teachers' Tools for the 21st Century: A Report on Teachers' Use of Technology*, approximately one-third of teachers reported feeling well prepared or very well prepared to use computers and the Internet for classroom instruction (National Center for Education Statistics, 2000).

Reasons for the resistance to adopt technology are numerous. However, the major barriers to implementation among teachers have been identified as a lack of time (Maney, 1999) and lack of support (McKenzie, 1998) and are discussed below.

Significance of this Study

According to a report from the research of the CEO Forum on Education and Technology (2000), while the majority of schools in the United States have access to technology, few teachers are ready to use technology in their lessons. The National Center for Educational Statistics (NCES) reports that nearly all public school teachers (99 percent) reported having computers available somewhere in their schools in 1999; 84 percent had at least one computer available in their classrooms, and 95 percent had computers available elsewhere in their schools. A Department of Education survey found that approximately 33 percent of teachers believe that they could use technology in their classrooms (NCES, 2000). The CEO Forum on Education and Technology (2000) further states that on average, while schools are spending $88 per student on computers, a mere $6 is spent on teacher training.

Research suggests that certain conditions, such as time for learning and planning, administrative support, and staff development, do influence technology adoption among teachers; however, such a lengthy list is overwhelming and certainly impractical to implement. Schools, which continually deal with budget constraints, need to know which factors may be the most important. A key to this process is informing administrators and elected officials about the need and possible solutions supported by research.

If, in fact, effective technology use does enhance student learning (Apple Classrooms of Tomorrow, 1996), then appropriate incorporation of technology is essential. The teacher, however, is the heart of this issue.

If student achievement is the ultimate goal, and if technology can be an effective aid in the fulfillment of that goal, then it is logical to assume that teacher implementation of technology will assure student implementation of technology. Kirkpatrick and Cuban (1998) identify teaching as the key ingredient to any successful technology adoption. Helping teachers use technology effectively may be the most important step to assuring that current and future investments in technology are realized (OTA, 1995).

It is evident that while computers and related peripherals are increasingly present in schools, teachers are lagging behind when it comes to integrating these technologies (OTA, 1995). School administrators and technology coordinators are attempting to address the issue of teachers and technology integration, but they are guided mainly by the recommendations and suggestions of others. That is, these decision makers are relying solely on the advice of education experts. Such advice is often not based on empirical research. As has been stated, a mere 20-30 percent of teachers believe they could integrate some kind of technology into their classrooms (NCES, 2000). Maney (1999) and Sandholtz, Ringstaff, and Dwyer (1997), identify the teacher as the key to the integration of technology in the classroom. However, teacher issues often are overlooked (McKenzie, 1999). While research suggests that conditions exist which do influence teacher use of computer technology, it is not known which conditions or combination of conditions are more likely to influence or be related to the secondary teacher's decision to adopt selected educational technology in the curriculum.

The Contributions of Technology to Education

Before technology will be adopted into the educational process, teachers must first be convinced that there is some advantage to using technology. Findings from the Apple Classrooms of Tomorrow (ACOT) research suggest that students' behaviors and attendance improved, along with attitude toward themselves and toward learning (Ringstaff, Yocam, & Marsh, 1996). Improved student performance was also noted in several ways: (1) test scores indicated that, despite time spent learning to use technology, students were performing well, and some were clearly performing better, (2) the students wrote more, more effectively, and with greater fluidity, and (3) some classes finished whole units of study far more quickly than in past years (Ringstaff, et al, 1996). Other unintended outcomes were noted which included students becoming socially more aware and more confident, students working well collaboratively, and students exploring and representing information dynamically and in many forms (Ringstaff, et al, 1996).

Educational technology makes it possible to create learning situations in which students can be engaged in activities that they find interesting and exciting for their own reasons and which accomplish the educational goals of teachers (Riel, 1989). Teachers can plan various activities simultaneously, and students can learn in an interactive, workshop-style format (Piper, 2000). Computers give teachers the opportunity to expand the boundaries of the classroom (Facemyer & Peterson, 1996, as cited in Piper, 2000) by allowing instant access to information and specialized expertise. In addition, educational technology can create new avenues for social exchange and cooperative learning (Riel, 1989).

Pre-service teacher training. Few school districts and accrediting organizations are currently requiring technology proficiency, although that is rapidly changing. In a recent survey of teachers, the majority of new teachers recalled that technology proficiency was not required to obtain teaching certification (67.3%) or their first teaching position (82.9%). Further, teacher colleges are evidently not doing their part in educating future teachers in integrating technology into the classroom. While the majority of teacher education institutions report that their technology infrastructure is at least adequate to carry out their current programs, about one third still believe deficiencies in their facilities limit their programs (CEO Forum, 1999).

An estimated 40percent of present teachers will retire or leave the profession by the year 2004. Yet slightly more than one-third of new teachers reported that their college experience left them "very well prepared" or "well prepared" to integrate technology into classroom instruction (MDR, 1999). Further, more than 70 percent of teacher preparation programs require three or more credit hours of instruction in courses focused on technology. About half of that instruction is included in other classes such as methods and curriculum courses. Twenty-five states presently require "computer education" for initial licensure. The National Council for the Accreditation of Teacher Education (NCATE) has promoted the preparation pre-service teachers with educational technology and is addressing this issue. NCATE has released a series of technology-related curriculum guidelines that schools of education must meet before they receive accreditation with the goal that new teacher licensure programs should require proficiency in integrating technology into the curriculum by 2003 (CEO Forum, 1999). However, these standards apply only to those colleges and universities that seek NCATE accreditation.

Methodology

This descriptive survey research determined the extent to which certain variables are related to a teacher's decision to adopt selected technology. Descriptive research was appropriate for this study because relationships and effects among variables were studied as they occurred in a natural setting (Wiersma, 1995).

Participants

This study was a type of non-random, purposeful sample called a comprehensive or census study, which used a homogeneous sampling design limited to secondary teachers at a northwest Ohio suburban school district. There were approximately 200 secondary teachers under contract in this district. The faculty consisted of approximately 80 males and 120 females, teaching in 11 different content areas. A suburban district was chosen because suburban districts represent "middle America" and is considered "typical case" which represents the largest population in the United States. A census study was appropriate for typical case or "middle of the road" sampling (Wiersma, 1999). The results of a census study, however, are not appropriate for generalizing to a larger population since the sample is non-random. However, a logical basis for generalization may be argued if populations from the school district are comparable with nationwide.

Instrumentation

The background information obtained from the subjects on the survey included gender, age, content area taught, highest college level attained, years of teaching experience, the year certification was granted, possession and use of a home computer, computer experience, and current stage of technology adoption. This data served as a descriptive profile of the subjects involved in teaching secondary students under ordinary conditions. The purpose of this data was also to increase the possible external validity of the results.

Four parts comprised the remainder of the survey. Part 1 dealt with staff development items, Part 2 dealt with the relationship of physical resources and effective technology integration, Part 3 dealt with attitude toward technology, and Part 4 was designed to obtain baseline data of teacher skill and confidence levels.

A reliability analysis of the survey ***Teachers & Technology Survey: Assessing the Present, Planning for the Future*** used in this research was conducted. Inter-item reliability, as measured by Cronbach's Alpha, was estimated at .93. Therefore, the reliability of this survey is high, meaning that the survey should yield the same, or nearly the same, ranks over repeated administrations. Of course, it must be mentioned that this assertion only holds true if the trait being measured has not changed.

Administration of Research Instrument

The survey instrument was administered to all secondary teachers in a typical suburban school district in Northwest Ohio during a district wide inservice meeting. All teachers from all content areas were asked to fill out all sections of the survey dealing with staff development, physical resources, attitude toward technology, base-line data on skill and confidence levels, and demographics.

Method of Analysis

For this descriptive, qualitative survey study, descriptive statistics were appropriate to analyze data. Inferential statistics do not apply to a census, because population measures are parameters and it is not appropriate to test hypotheses about them or to estimate them.

Correlations and correlation coefficients, using the Pearson Product Moment Formula, were calculated to describe the relationship between: (1) staff development and technology adoption, (2) physical resources and technology adoption, and (3) attitude toward technology and technology adoption. The Pearson Product Moment was appropriate for the analysis because it required at least interval scale measurement (Wiersma, 1995). A correlation matrix was calculated by the statistical program SPSS. High, positive correlations would suggest a more positive attitude is related to a higher rate of technology adoption. High, negative correlations would indicate a reverse of those relationships. Lower correlations would suggest little to no relationship between these variables. Theoretically, an interval scale is needed to compute means and correlations and technically, Likert scales are ordinal; however, if there are a number of items (usually at least 20), with a similar number of items in the subscales, an interval scale is approached for statistical purposes (Wiersma, 1995). The results from this data were calculated using the SPSS statistical package.

Finally, the reliability of the instrument was estimated using the SPSS statistical package. Determining reliability was important to the interpretation of the results. The higher the reliability, the more confident we can be that such results would be reproducible in other situations, under similar conditions. Individual items were further correlated to strengthen the validity of the instrument.

The following correlation coefficient scale is used for reporting and interpreting results.

Table 1. *Correlation Coefficient Values*

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to –1.00)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (-.00 to -.30)	Little, if any, correlation

Note. From Applied statistics for the behavioral sciences p. 120, by D. E. Hinkle, W. Wiersma, and S. G. Jurs, New York: Houghton Mifflin. Copyright 1998 by Houghton Mifflin Company. Adapted with permission.

Items from each section on the survey were grouped and summarized by category. Items from Part One of the survey have been collapsed into a category labeled Staff Development. Items from Part Two have been collapsed into a category labeled Physical Resources. Finally, items from Part Three have been collapsed into a category labeled Attitude Toward Technology. Likert scales are ordinal; however, if there are a number of items (usually at least 20), with a similar number of items in the subscales, an interval scale is approached for statistical purposes (Wiersma, 1995). In addition, mean scores, standard deviations and the correlation to technology adoption are reported for each individual item within sub-categories. The results from these data have been calculated using the SPSS statistical package.

The main categories of Staff Development, Physical Resources, and Attitude Toward Technology have been correlated with the variable technology adoption and will be discussed below. The results are presented in Table 2. Moreover, the category Current Rate of Technology Adoption was correlated to the category Stage of Technology Adoption. Not only was the correlation significant, the categories showed a moderately high correlation (see Table 2).

Table 2. *Inter-correlations Between Subscales With Current Rate of Technology Adoption*

Subscale	Rate of Technology Adoption	Stage of Technology Adoption	Staff Development	Physical Resources	Attitude Toward Technology
		Teachers	(N = 132)		
Rate of Technology Adoption	--	.646**	.295**	-.172*	-.044
Stage of Technology Adoption		--	.280**	-.172*	-.074
Staff Development			--	.382**	.031
Physical Resources				--	.081
Attitude Toward Technology					--

* $p < .05$ **$p < .01$

Implications

Given the results of this study concerning the conditions which facilitate the adoption of technology among secondary teachers, it is in the judgment of this author that there are several implications that may be offered in the area of technology adoption and the change process in the K-12 environment. It is important to keep in mind that this study was a census study, and results can be generalized to other populations on a logical basis only.

Certain demographic variables seem to be related to technology adoption. These variables have been cited in various studies and include younger, male teachers as the likely technology adopters. These variables are, for the most part, unalterable. However, younger teachers are more likely to adopt because of greater exposure to technology in college classes and in life in general.

If administrators and others stakeholders want to encourage technology use, then certain conditions need to be created which promote the use of the technology. The list of conditions, which follow, was determined not only by the rating of the participants, but by the strength of correlations of each variable with technology adoption. These will be listed in order of importance.

Provide Staff Development Opportunities. The most important element, according to this research, was staff development. Staff development is critical to the adoption process. This study indicated that formal classes and the "trial and error" approach are the preferred methods for learning to use the computer. School administrators can promote technology integration by providing numerous staff development opportunities for their teachers. These can take the form of inservices, conferences, and classes at local colleges/universities paid for by the school district. Training activities can also be less formal, in which teachers work together to learn new technologies, and trial and error approaches are encouraged.

Provide Time. Teachers rated time as the second most important factor promoting technology adoption. Time is needed to learn new technologies and time is needed to learn to integrate technologies in the classroom. When not enough time is provided, teachers often resort to what is already known, and so fail to try new methods to teach students. If technology adoption is expected, then the time must be made available for learning the technology, peer tutoring, and cooperative ventures.

Provide Prompt Technical Assistance. Third, prompt technical assistance was rated highly by the teachers. If equipment is not going to work, and if no one is available to help when the inevitable operational problems occur, then teachers will be less likely to use the equipment available. That is, teachers won't use technology if it is perceived to be a hassle. The need for prompt technical help appears to be an important variable, as indicated by level of agreement among the participants.

Provide Incentives. Two other variables are worth mentioning, though they do not appear to be as important as those mentioned. The need for good reasons, or incentives, to use new technologies was highly rated by the teachers. The incentives may be internal, such as providing reasons supported by empirical research for using technology. The incentives may also be external, such as allowing adopting teachers a greater voice in equipment recommendations or rewarding adopting teachers with the newer equipment.

Promote Positive Attitudes. This study supports previous research that certain attitudes are important to the process of technology adoption. In general, teachers need to see the effectiveness of technology on student learning outcomes. If teachers are allowed to share in the decision-making process, then technology adoption is likely to occur. Increased confidence in using technology and more positive attitudes toward technology can be promoted by increasing the exposure teachers have with the technology. This can usually be accomplished through training and staff development activities, allotted time, and incentives proved to use the computers.

Recommendations for Further Study

This study was purely descriptive in nature. Correlations were reported, and some were found to be significant. However, correlation does not imply cause and effect. Nor does correlation predict the effect of one variable on the attainment of another variable. Correlation only indicates the strength significance of the association of variables. Therefore, further studies should be conducted in experimental research which examines more closely the relationship between variables. The purpose of such research would be to determine if these relationships, suggested by correlation, indicate prediction in the adoption of technology. That is, research would then become more prescriptive in nature. A set of proven guidelines or criteria could be established which administrators would use to promote

Time and staff development have been frequently mentioned in the literature as two of the most important factors for technology adoption. This study agreed with the research. However, time is a very vague variable. Future studies should concentrate on the amount of extra time needed and how that time should be scheduled. That is, do teachers merely need extra time, or should the time be structured in some way? If the time is structured, what is the best structure? Options that should be studied include time used in formal training time, inservice time, time used for peer tutoring or coaching, time spent with a technology specialist, time spent with a student technology specialist, and so on. Moreover, time in the student schedule should be addressed. The optimal time for student learning with technology is not yet known.

In summary, the value of this research should be the practicality to administrators and principals interested in the process of technology adoption. If change is to occur in the classrooms, it must begin with the teacher. Research has suggested that conditions exist which do influence teacher use of computer technology. This study attempted to answer the question which conditions or combination of conditions are *more likely* to influence or be related to the secondary teacher's decision to adopt selected educational technology in the curriculum. The answers, as determined in this research, came from those who can most successfully effect change in the classroom – the teachers.

References

Apple Classrooms of Tomorrow. (1996). Changing the conversation about teaching, learning & technology. A report on 10 years of ACOT research. [online] 18 pages. Available: http://www.apple.com/education/k12/leadership/acot/library.html

Becker, H. J. (1999). Internet use by teachers: Conditions of professional use and teacher-directed student use. Teaching, Learning and Computing: 1998 National Survey, Report #1. [online] 31 pages. Available: http://www.crito.uci.edu/TLC.findings/Internet-use/startpage.htm

CEO Forum on Education and Technology. (1999). Professional development: A link to better learning. The CEO Forum School Technology and Readiness Report, Year Two. [online] 41 pages. Available: http://www.ceoforum.org

CEO Forum on Education and Technology. (2000). The power of digital learning: Integrating digital content. The CEO Forum School Technology and Readiness Report, Year Three. [online] 32 pages. Available: http://www.ceoforum.org

Firek, H. (April 2003). One order of ed tech coming up… You want fries with that? Phi Delta Kappan 84 (8), 596-597.

Hinkle D. E., Wiersma W., & Jurs, S. G. (1998). Applied statistics for the behavioral sciences. New York: Houghton Mifflin.

Kirkpatrick, H. & Cuban, L. (1998).Computers make kids smarter – right? Technos Quarterly for Education and Technology 7 (2). [online] 9 pages. Available: http://www.technos.net/journal/volume7/2cuban.htm

Maney, J. K. (1999). The role of technology in education: Reality, pitfalls, and potential. In G. Cizek (Ed.), The handbook of educational policy (pp. 387-415). San Diego, CA: The Academic Press

Market Data Retrieval. (1999). MDR releases yearly school technology report. [online] 2 pages. Available: http://www.schooldata.com/pr18.html

McKenzie, J. (March 1999). Secrets of Success: The ten most important staff development lessons. Session presented at: Michigan Association for Computer-related technology Users in Learning Conference, Detroit, MI.

McKenzie, J. (1999). How teachers learn technology best. Bellingham, WA: FNO Press.

McKenzie, J. (1999). The technology in my life survey. From Now On The Educational Technology Journal [online]. Available: http://emifyes.iserver.neet/fromnow/techlife.html

McKenzie, J. (1998). Secrets of success: Professional development that works. eSchool News [online] 11 pages. Available:http://www.fromnowon.org/eschool/secrets.html

Piper, D. M. (2000). Facilitating change: Factors influencing teachers' practices of using computers in the classroom. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Indiana, PA.

Riel, M. (1989) The impact of computers in classrooms. Journal of Research on Computing in Education, 27 (2), 180-190.

Ringstaff, C., Yocam, K. & Marsh, J. (1996). Integrating technology into classroom instruction: An assessment of the impact of the ACOT teacher development center project. Apple Classrooms of Tomorrow [online] 18 pages. Available: http://www.apple.com/education/k12/leadership/acot/library.html

Rogers, E. M. (1995). Diffusions of innovations (4th ed.). New York, NY: The Free Press.

Sandholtz, J. H., Ringstaff, C., & Dwyer, D. C. (1997). Teaching with technology: Creating student-centered classrooms. New York: Teachers College Press.

U.S. Congress, Office of Technology Assessment. (1995) Teachers and technology: Making the connection. Washington, D.C: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics. (1999). Digest of Education Statistics, 1998. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics. (2000). Teachers' tools for the 21st century: A report on teachers' use of technology. Washington, D.C.: U.S. Government Printing Office.

Wiersma, W. (1995). Research methods in education. Needham Heights, MA: Allyn and Bacon.

The Lee and Lida Cochran AECT Archives: Who, When, Where, Why, and How?

Rebecca P. Butler
Judith Rabak-Wagener
Northern Illinois University

Introduction

Today, the Lee and Lida Cochran AECT Archives reside in the College of Education at Northern Illinois University in DeKalb, Illinois. This grouping of audio and visual equipment from the early days of instructional technology form a vibrant part of the college's Blackwell Museum of Education and represents a unique collection of hardware, which supports research into the early days of the field. Why were the archives created? How was this accomplished? When did this happen? Where have they been stored? Who are Lee and Lida Cochran? These questions and more are answered in the course of this article.

Why Were the AECT Archives Created?

The history of the archives begins in 1952 when a National Education Association (NEA), Department of Audio-Visual Instruction (DAVI) committee meeting, headed by Dr. E. Winifred Crawford, turned to the topic of audiovisual education history. At that time DAVI (an organization founded in 1923 under the umbrella of the NEA) (1) was concerned that the history of the new field of audiovisual education might be lost. Discussion eventually turned to reality in the mid-fifties, with the placement of the DAVI Archives in the library at the University of Iowa. (Boone, 1998)

What Is the History of The AECT Archives?

Once housed at the University of Iowa, the archives began to grow. From a collection of DAVI proceedings and minutes of past meetings, the archives expanded (print) to books, pamphlets, correspondence, and records and (non-print) to historical audiovisual equipment. These machines would eventually range from early visual devices such as a zoetrope and a stereopticon (2) to benchmark film projectors and audio recorders. (3) In addition, Rita Hochheimer, one of the first women in the field, "…gave nearly 200 books, pamphlets, and bulletins from her personal files" (Boone, 1998, 24). Over the next few years, a variety of individuals, manufacturers, organizations, and institutions of K-12 and higher learning from across the nation also provided print and equipment donations.

By the early sixties, the interest in preserving educational media had moved beyond DAVI. Thus, when a movement for a combined motion picture/audiovisual museum began, the archives were shipped to Hollywood, California, where they were placed in storage until the proposed museum could be constructed. When the museum did not materialize as planned, the hardware portion of the archives was moved to the Los Angeles County Museum of Natural History, while the rest of the collection remained in storage. By 1968, however, both the hardware and print software portions of the archives were back at the University of Iowa in Iowa City. There they remained until 1994, when the print and software part of the archives moved to the National Public Broadcasting Archives at the University of Maryland's Hornbake Library. (Boone, 1998; Butler, 1998) (4) Continued discussion of the archives below, focuses on the hardware or equipment portion of the DAVI Archives, which would eventually become the AECT Archives and then the Lee and Lida Cochran AECT Archives.

While the print and software portion of the archives traveled to Maryland, the hardware portion of the collection remained in storage at Iowa until 1995, when it moved to its new home at Utah State University, Logan, Utah. At Utah State, the hardware archives were housed in a new science and technology building. There it remained until December 2000, at which time a young Utah State couple spent their honeymoon driving it out to DeKalb, Illinois, in a rental truck. Once there, volunteers spent a day unloading the equipment into the basement of the Milan Township One Room School on the campus of Northern Illinois University. Today the Archives, used for research and display purposes, help support a doctoral seminar on the foundations of the field of instructional technology in the Department of Educational Technology, Research and Assessment in the College of Education at Northern Illinois University. Various pieces of the hardware are also displayed in the college's Learning Center, and plans are to highlight the collection with conferences, presentations, and displays. A web site for the Archives is available at < **http://www.cedu.niu.edu/blackwell/multimedia/** >.

Who Are Lee and Lida Cochran?

The archives have gone through several name changes. They started out as the DAVI Archives in the early 1950s. When the name of the parent organization changed to AECT in 1970, the name of the archives also changed – to the AECT Archives. The addition of Lee and Lida Cochran's name occurred after the archives moved to Northern Illinois University. So – who exactly are Lee and Lida Cochran and why do their names grace the AECT Archives today?

Lee Cochran joined the audiovisual education field in 1923, when, as a high school student he began working for the Extension Division at the University of Iowa, Iowa City, Iowa. He stayed working at the University of Iowa for most of his life. In 1964, he became the Director of the Audio-Visual Center there, a position from which he retired in 1969. Lee was a self-taught media person. Throughout five plus decades, his interest and involvement in the field led to a variety of positions within the important audiovisual organizations of the time. For instance, he was president of DAVI from 1954-1955, and the primary leader in the establishment of the Lake Okoboji Educational Media Leadership Conferences (5) which took place at Lake Okoboji, Iowa, from 1955 to 1970. In addition, Lee was instrumental in bringing the AECT Archives to the University of Iowa in the early fifties and active in their return to Iowa after their sojourn in Hollywood. He was prominent in building the collection throughout its tenure at Iowa, as well. Because of his decades long influence on the collection, it is only fitting that his name is one of two attached to the archives.

So, why is Lida Cochran's name also attached to the AECT Archives? Lida, too, is very prominent in our field. She started out in the 1950s in the state of Washington as a photographer. Through her interests and skills in media, within a few years she was teaching at Central Washington State College and it was this job that would eventually place her in Iowa. "…the college sent me back through the Midwest to look at the various audiovisual services to get ideas, because I was to do this at Central Washington. The first place I came to was Iowa, because Lee (Cochran) had developed a program here that was recognized as being a very good one and …so that is when I met Lee Cochran" (Cochran, 1992). Lida married Lee in 1961 and began teaching audiovisual courses at the University of Iowa soon after that. "I taught selection on utilization of audiovisual materials, and then I taught production… one semester I taught non-projected materials and then the next semester I taught projected materials…" (Cochran, 1992). However, Lida's fame in our field stems from her enthusiasm and activism in visual literacy. She is an early member of the International Visual Literacy Association (IVLA) and a "mover and shaker" in this aspect of instructional technology. "The concept of visual literacy is very broad and involves every discipline there is. …to be a visual literate, we feel that one also has to be able to create pictures whether you draw them or if you use a media camera or a still camera. …visuals are a language" (Cochran, 1992) (6)

Thus, it is Lee's influence on the collection and audiovisual field and Lida's prominence in visual literacy that make the addition of their names to the AECT Archives appropriate.

What Types of Equipment Now Reside in the Lee and Lida Cochran AECT Archives?

The hardware currently in the archives numbers in the hundreds. It ranges from the 1800s (early precursors to the film projector) to "antique" computers of the 1980s. A few of the pieces are briefly described below.

The Polyorama Panoptique

The Polyorama Panoptique dates from the 1859s and is an optical toy. Using French and English landscape images, landmarks, and people, the archives' version is a small hand-held device. Using a main box of oak, an adjustable lens, a set of paper images mounted on wooden frames, and a hinged door, light is let into the back of this piece. Users view the image through a lens in the back door. Action occurs and scenes change as the light is varied (Wiegmann, 2003).

The Magic Lantern

The Lee and Lida Cochran AECT Archives is home to several magic lanterns of a variety of sizes. Most date from the 1800s. Magic lanterns probably started out for home use, but they soon became popular in the schools as well. In essence, the magic lantern is a forerunner of the slide projector. Images painted on glass slides are "…inserted between a light source and a series of objective lenses, which projected the image onto walls, screens, or cloth drapes." (Stepien, 2003, 3) Magic lanterns were often powered by kerosene. Small toy ones might also use a candle, and the archives have one which is powered by whale oil.

The Victor Animatograph Cine` Projector

Dated 1923, this hand-cranked motion picture projector also has an electric motor. It can be operated by either method and runs 16mm film (Clemens, 2003).

The Magnetic Wire Recorder

The archives' model is a 1948 RCA sound recorder. Developed to supplement radio equipment of the time, the wire recorder – where sound was recorded on wire looped into a metal cassette -- was considered a cheap and portable recording and reproduction mechanism. It is considered the forerunner of cassette audio and video recorders (Bushong, 2003).

The TELEX 610 Series Headphones

One of the archives' "newer" pieces of hardware, these headphones date from 1963 and originally cost $6.30. Mainly sold on the East coast to K-12 schools, these headphones are so durable that some are still in use today (Lausch, 2001).

Much fascinating equipment exists in the archives and more pieces are added as they are donated to the collection.

Conclusion

The Lee and Lida Cochran AECT Archives, a compilation of early audiovisual hardware, are an example of benchmark equipment in the field. Housed at Northern Illinois University, this collection exists as a representative of the history of the field of audiovisual education/instructional technology: research, teaching, demonstration, display, and reflection.

End Notes

(1) In 1970, DAVI moved out from under the umbrella of the NEA to become the Association of Educational Communications and Technology. (Settler, 1990)

(2) These are discussed later in this paper.

(3) Today, the audiovisual pieces in the archives also include cameras, early computers, slide projectors, and a variety of recording and playback devices, among other types of equipment.

(4) The AECT Print Archives are located in the National Public Broadcasting Archives at Hornbake Library, University of Maryland, College Park, Maryland. The print archives focus on the beginning days of visual/audiovisual education/instructional technology in the United States. Material in the AECT Print Archives includes membership listings, convention proposals, committee reports from a variety of early audiovisual organizations, teaching guides and manuals, film and audiovisual catalogs, and private correspondence from early audiovisual and instructional technology pioneers. The print archives also contain a number of oral history interviews, both audio tapes and transcriptions, of several of the founders of the field (Butler, 1998).

(5) The Lake Okoboji Educational Media Leadership Conferences, established as a national forum to discuss the problems and issues of the audiovisual field, were comprised of members of DAVI nominated by the organization. Forty-five attended the first year (Penn, 1972).

(6) Lida, whose name is almost synonymous with visual literacy, was rejected as a presenter at the first IVLA. She says that her program idea, entitled "The Fine Art of Seeing" "…wasn't close enough to the top and they had some brilliant people at that first conference" (Cochran, 1992).

References

Boone, Luke L. 1998. "Our AECT Archives." *TechTrends* 43, no. 1 (January/February): 23-26.

Bushong, Rich. 2003. "1948 RCA M-1 (Magnetic Sound) Wire Recorder." Unpublished paper. DeKalb, IL: Northern Illinois University.

Butler, Rebecca P. 1998. "Preserving the History of Our Field: The AECT Archives." *TechTrends* 43, no. 1 (January/February): 27-32.

Clemens, Alan. 2003. "The Victor Animatograph Cine` Projector." Unpublished paper. DeKalb, IL: Northern Illinois University.

Cochran, Lida. 1992. Interview with Rebecca P. Butler.

Lausch, Mark. 2001. "AECT Archive Research: TELEX 610 Series Instructional Headphones." Unpublished paper. DeKalb, IL: Northern Illinois University.

Penn, Vivian Sue. 1972. "A Study of the Origin, Structure and Significance of the Lake Okoboji.Educational Media Leadership Conferences, 1955-1970." Unpublished dissertation. Philadelphia: Temple University.

Saettler, Paul. 1990. The Evolution of American Educational Technology. Englewood, CO: Libraries Unlimited.

Stepien, Bill. 2003. "AECT Archive Project: The Magic Lantern." Unpublished paper. DeKalb, IL: Northern Illinois University.

Wiegmann, Beth. 2003. "AECT Archive: The Polyorama Panoptique." Unpublished paper. DeKalb, IL: Northern Illinois University.

Personality as a Grouping Strategy for Online Collaborative Learning

Edward J. Caropreso
Sue-Jen Chen
University of North Carolina at Wilmington

Abstract

Learning occurs within a context of communication, including online collaborative discussions. This study investigated the effects of personality on online communication, task engagement, and students' attitude toward online discussion. The results indicated that: 1) personality influences the types and patterns of communication; 2) level of task engagement; and 3) students' feelings, attitudes and opinions about online discussions are positive. To optimize the effect of online discussion and collaboration, grouping strategies based on personality should be considered.

Introduction

Learning involves communication in many forms. Human communication has historically been interpreted in a variety of ways, including communication as both cultural and social processes. As a cultural process, communication involves learning and using language conventions such as postures, gestures, symbols and their arrangements that have shared or agreed-upon interpretations. As a social process, communication becomes the principal way in which human beings experience meaningful interactions. Through such meaningful interactions, people learn to play roles, understand social norms, recognize and apply social sanctions and evaluate each other's actions according to systems of shared values and beliefs (DeFleur & Bal-Rokeach, 1982). Clearly, many of these features of human communication are unavailable or ambiguous during online discussions. Despite the clear importance and relevance of communication as engaged cultural and social interaction, a question remains regarding how communication during online learning experiences may be influenced by such absences or ambiguities.

The value of collaborative learning experiences has been well established (Johnson, & Johnson, 1989; Johnson, Johnson, & Smith, 1997). Educators believe students benefit from collborative learning, and Internet technology provides important opportunities for various kinds of collaborative learning via synchronous and asychronous communication channels. Collaboration has been a growing component of online learning research, increasingly attracting attention and resources.

Online collaboration has been enthusiatically studied by many educators over the past decade. Previous research has focused several aspects of online learning, including types of collaborative strategies, community building, instructional design, and learning styles. Some commonly used grouping strategies, such as random assignment and self-selection into groups, have been problematic for many students when dealing with team-based tasks. Thus, creating and sustaining effective learning communities remain challenging for the instructors. Effective grouping and communication strategies appear to be the essential parts of the online learning process (Sherry, 1996; Sherry, Billig, & Tavalin, n.d.).

Online collaboration has typically been studied within the context of learning communities. Though communities are composed of groups of people with common interests (American Heritage Dictionary, 1997), individuals within communities bring their own unique personalities. Little attention has been paid to the potential influence of students' personality on collaborative group discussion. Personality is a well-known individual differences variable, a stable trait, on which people vary consistently. In other words, an influence of personality is likely to be a consistent feature of group-based interactions. Personality influences attitudes, beliefs and behaviors (Saucier & Goldberg, 2003); therefore, an individual's personality will likely influence inter- and intra-personal interactions with members of a community. When collaborative activity occurs in the context of learning experiences, approppriate pedogogical strategies must be identified and implemented (Comeaux, 2002, p. xviii). To optimize the effectiveness of a learning community, the potential influence of personality must be taken into consideration as part of the design of online collaborative learning activities.

Purpose of the Study

1. How does personality influence group interaction and type and pattern of online communication?
2. How does personality influence task engagement during online discussion forums?
3. What are students' feelings, attitudes and opinions about their online discussion experiences?

Methodology

Participants

Seventy-three undergraduate students enrolled in three sections of an undergraduate educational psychology course participated in the study. Based on the results of the International Personality Item Pool (IPIP) online personality inventory (Buchanan, 2001; Buchanan et al., 1999; Buchanan & Smith, 1999; Goldberg, 1999b) completed prior to the start of the online discussions, students were assigned to different personality profile groups. Three traits were identified that theoretically might influence engagement and participation in online collaborations: Extraversion; Agreeableness; and Openness. Extraversion reflects a tendency to seek and engage in social interactions; agreeableness reflects interaction quality; and openness reflects an interest in intellectual and imaginative experiences. High scores relative to the scores of the entire sample indicted a greater tendencies toward these traits. Students were identified as "relatively high" if they scored at or above the 67th percentile and "relatively low" if they scored at or below the 33rd percentile on these three traits. All other students were considered as "neutral;" these students' data were not analyzed in this study. Our particular interest was in the potential differences in communication and task engagement for the extremes in measured personality, therefore, we only included "High," "Low" and Combined "High and Low" groups in our analyses. Across the three sections, we identified a "High" group of n = 16, a "Low" group of n = 9 and a "Combined" group n = 9; the total sample size was N = 34. The sample sizes varied across the three discussions due to variations in participation rate and class attrition.

Procedure

Students were assigned to participate in three online discussions about three different case studies, scenarios that reflected important course concepts. Discussions occurred at three different times, about once per month, from the beginning till the end of the semester. Open-ended focus questions were provided to guide the online discussions; in addition, the general purpose was to discuss the scenarios using relevant course concepts and content, as each scenario reflected specific course material and outcomes being addressed in class during each online discussion. WebCT online course management system was used as the forum for the online discussions. Each discussion had specific openning and closing dates, staying open from 7-12 days depending on the nature of the case study and concepts under consideration. Each group member was required to login and post a minimum of three messages during the discussion forum. Postings were assessed for course credit; each posted message was worth a specific value up to a possible maximum of nine posted messages across all three discussions. No credit was earned for additional postings. After the third discussion, students completed an inclass survey regarding their feelings, attitudes, and opinions about using online discussion and collaborative teamwork to enhance learning. Student responses were measured using a 6-point Likert scale, with 6 representing "strongly agree" and 1 "strongly disagree."

Data and Analyses

A total of nine messages (three for each discussion) for each student in the High, Low or Combined personality groups were analyzed for the pattern and type (one-way vs. two-way) of communication and degree of task engagement (fully engaged; somewhat engaged; disengaged). One-way communication (scored 1) involved messages expressing questions, comments, statements or opinions, but neither inviting, encouraging, nor soliciting reactions from group members. Two-way communication (scored 2) involved inquiry and/or responding messages engaging other members through questioning, commentary, statements or opinions explicitly responding to previous messages or directly inviting, encouraging or soliciting replies.

Task engagement was interpreted as the degree to which messages related the case studies and course concepts. Being "Fully engaged" (scored 3) involved content that specifically and consistently focused on instructional or assignment issues. "Somewhat engaged" (scored 2) was used to code content that clearly but inconsistently reflected instructional or assignment issues, and "Disengaged" (scored 1) reflected content that was either marginally related or unrelated to instructional or assignment issues. Percent of agreement of ratings across two trained raters was used as the measure of interrater reliability. Each rater independently coded all messages from Discussion 1. Codings for each message were compared for degree of agreement; disagreements were resolved through discussion and all messages were recoded. The final level of agreement for communication type was 94% and for task engagement was 91%.

In order to answer the research questions, descriptive analyses of the data for communication type, task engagement, and survey responses were conducted. Cross tabulations were used to compare the percentage of communication type, reflecting communication patterns, and engagement level of messages within each discussion posted by the High, Low and Combined personality groups.

Results and Discussion

Tables 1 through 3 show within group results for communication type. Within-group comparisons involved looking at the percent of each group's communication type for each message. These within-group results were the basis for identifying communication pattens that profiled each group. For every message across all three discussions, the High group consistently exceeded both the Low and Combined groups in the percent of two-way communication messages. For example, in 8 of 9 messages across three discussions more than 50% of the High group used two-way communications, while only 2 of 9 of the Low group and 5 of 9 of the Combined group used two-way communication. In very few instances, is this pattern not demonstrated. For example, in only two cases did the Low group exceed the High group, though in both cases the High group demonstrated more than 50% engaged in two-way communication. In only three cases did the Combined group's percentage of two-way communincation exceed the High group's, but in two of these, the High group again exceeded 50% or more engaged in two-way communication.

Table 1: *Within-group percentage comparison of communication type for discussion one*

Discussion	1								
Message	M1			M2			M3		
N	33			32			29		
Groups	H	C	L	H	C	L	H	C	L
n	15	9	9	15	9	8	13	8	8
One-way	33.3	100	66.7	33.3	44.4	50.0	46.2	62.5	37.5
Two-way	66.7	0	33.3	66.7	55.6	50.0	53.8	37.5	62.5
Total	100	100	100	100	100	100	100	100	100

Note: H=High personality group; C=Combined High and low personality group; L=Low personality group

Table 2: *Within-group percentage comparison of communication types for discussion two*

Discussion	2								
Message	M1			M2			M3		
N	34			31			30		
Groups	H	C	L	H	C	L	H	C	L
n	16	9	9	14	9	8	14	8	8
One-way	62.5	44.4	100	21.4	33.3	75.0	42.9	50	100
Two-way	37.5	55.6	0	78.6	66.7	25.0	57.1	50	0
Total	100	100	100	100	100	100	100	100	100

Note: H=High personality group; C=Combined high and low personality group; L=Low personality group

Table 3: *Within-group percentage comparison of communication types for discussion three*

Discussion	3								
Message	M1			M2			M3		
N	32			31			28		
Groups	H	C	L	H	C	L	H	C	L
n	15	9	8	14	9	8	12	8	8
One-way	46.7	44.4	75.0	42.9	33.3	37.5	33.3	50	75.0
Two-way	53.3	55.6	25.0	57.1	66.7	62.5	66.7	50	25.0
Total	100	100	100	100	100	100	100	100	100

Note: H=High personality group; C=Combined high and low personality group; L=Low personality group

Tables 4 through 6 show the cross-group results for each message for each discussion. Cross-group comparisons involved combining all three groups as the basis for comparing each group's percent of communication type. The results in Tables 4-6 parallel those presented in Table 1 through 3. Generally, the High group consistently exceeded both the Combined and Low groups in two-way communication. These analysis demonstrates that Highs exceeded all other groups across all messages in every discussion.

Table 4: *Cross-group percentage comparison of communication types for discussions one*

Discussion	1								
Message	M1			M2			M3		
N	33			32			29		
Group	H	C	L	H	C	L	H	C	L
One-way	15.2	18.2	27.3	15.6	12.5	12.5	20.7	10.3	17.2
Two-way	30.3	9.1	0	31.3	12.5	15.6	24.1	17.1	10.3
Total	100			100			100		

Note: H=High personality group; C=Combined high and low personality group; L=Low personality group

Table 5: *Cross-group percentage comparison of communication types for discussions two*

Discussion	2

Message	M1			M2			M3		
N	34			31			30		
Group	H	C	L	H	C	L	H	C	L
One-way	29.4	26.5	11.8	9.7	19.4	9.7	20.0	26.7	13.3
Two-way	17.6	0	14.7	35.5	6.4	19.4	26.6	0	13.3
Total	100			100			100		

Note: H=High personality group; C=Combined high and low personality group; L=Low personality group

Table 6: *Cross-group percentage comparison of communication types for discussions three*

Discussion	3								
Message	M1			M2			M3		
N	33			32			30		
Group	H	C	L	H	C	L	H	C	L
One-way	21.2	18.2	12.1	18.8	9.4	12.5	13.3	20	16.7
Two-way	24.2	6.1	18.2	25	15.6	18.8	26.7	6.7	16.7
Total	100			100			100		

Note: H=High personality group; C=Combined high and low personality group; L=Low personality group

Personality, interpreted as degree of extraversion, agreeableness, and openness, was associated with consistent patterns of online communication. As Tables 1 through 6 show, in terms of the percent of responses within and across groups across all three discussions, High personality students demonstrated proportionlly more group-based interpersonal engagement than both Combined and Low personality students. In other words, students with relatively higher levels of extraversion, agreeableness, and openness tended to engage each other more often and more consistently in two-way communication, seeking and encouraging interactive communication. Students in the Low personality group tended toward one-way communication.

In terms of task engagement, Tables 7 through 9 show that the Low personality group tended to have a higher percentage of messages scored 1, or disengaged, (6 of 9 messages), while the High personality group had lower percentage of such messages (3 of 9), and the Combined group had none scored 1. Tables 7-9 also indicate degree engagement; the High and Low personality groups appear to be somewhat similar in terms of messaages scored as 3, fully engaged, while the Combined group was consistently more fully engaged than both of the other groups.

Table 7: *Within-group comparison of learning task engagement for discussion one*

Discussion	1								
Message	M1			M2			M3		
N	33			32			29		
Group	H	C	L	H	C	L	H	C	L
n	15	9	9	15	9	8	13	8	8
1	0	0	11.1	6.7	0	12.5	7.7	0	25.0
2	46.7	11.1	11.1	53.3	33.3	25.0	53.8	37.5	25.0
3	53.3	88.9	77.8	40.0	66.7	62.5	38.5	62.5	50.0
Total	100	100	100	100	100	100	100	100	100

Note: 1 = Disengaged, 2 = Somewhat engaged, 3 = Fully engaged

Table 8: *Within-group comparison of learning task engagement for discussion two*

Discussion	2								
Message	M1			M2			M3		
N	34			31			30		
Group	H	C	L	H	C	L	H	C	L
n	16	9	9	14	9	8	14	8	8
1	0	0	33.3	0	0	62.5	0	0	25.0
2	31.3	0	33.3	57.1	0	25.0	35.7	0	62.5
3	68.8	100	33.3	42.9	100	12.5	64.3	100	12.5
Total	100	100	100	100	100	100	100	100	100

Note: 1 = Disengaged, 2 = Somewhat engaged, 3 = Fully engaged

Table 9: *Within-group comparison of learning task engagement for discussion three*

Discussion	3		
Message	M1	M2	M3

N	32			31			28		
Group	H	C	L	H	C	L	H	C	L
n	15	9	8	14	9	8	12	8	8
1	0	0	0	7.1	0	0	8.3	0	0
2	40.0	0	25.0	50.0	33.3	50.0	50.0	0	37.5
3	60.0	100	75.0	42.9	66.7	50.0	41.7	100	62.5
Total	100	100	100	100	100	100	100	100	100

Note: 1 = Disengaged, 2 = Somewhat engaged, 3 = Fully engaged

Tables 10 through 12 show the cross-group results, pooling all three groups for comparison. These results reveals that the Low personality group had a higher percentage of disengaged messages (scored 1), reflecting the pattern demonstrated in Tables 7-9. The tables also indicated that the Low personality group consistently had a lower percentage of fully engaged messages (9 of 9 comparisons). In general, the Combined personality group's pattern of task engagement paralleled that of the High group's.

Table 10: *Cross-group comparison of learning task engagement for discussion one*

Discussion	1								
Message	M1			M2			M3		
N	33			32			29		
Group	H	C	L	H	C	L	H	C	L
1	0	0	3.0	3.1	0	3.1	3.4	0	6.9
2	21.2	3.0	3.0	25.0	9.4	6.3	24.1	10.3	6.9
3	24.2	24.2	21.2	18.8	18.8	15.6	17.2	17.2	13.8
Total	100			100			100		

Note: 1 = Disengaged, 2 = Somewhat engaged, 3 = Fully engaged

Table 11: *Cross-group comparison of learning task engagement for discussion two*

Discussion	2								
Message	M1			M2			M3		
N	34			31			30		
Group	H	C	L	H	C	L	H	C	L
1	0	0	8.8	0	0	16.1	0	0	6.7
2	14.7	0	8.8	25.8	0	6.5	16.7	0	16.7
3	32.4	26.5	8.8	19.4	29	3.2	30.0	26.7	3.3
Total	100			100			100		

Note: 1 = Disengaged, 2 = Somewhat engaged, 3 = Fully engaged

Table 12: *Cross-group comparison of learning task engagement for discussion three*

Discussion	3								
Message	M1			M2			M3		
N	32			31			28		
Group	H	C	L	H	C	L	H	C	L
1	0	0	0	3.1	0	0	3.3	0	0
2	18.2	0	6.1	21.9	3.1	12.5	20.0	13.3	10.0
3	27.3	30.3	18.2	18.8	28.1	12.5	16.7	20.0	16.7
Total	100			100			100		

Note: 1 = Disengaged, 2 = Somewhat engaged, 3 = Fully engaged

Table 13 reports students' feelings, attitudes, and opinions about their online learning experiences, grouped into five categories. Regardless of personality profile, students were consistently positive and receptive.

Table 13: *Cross-group comparison of survey responses on feelings, attitudes, and opinions about online discussion*

Category	High (n=16)	Low (n=9)	Combined (n=10)	Total (N=35)
	M(SD)	M(SD)	M(SD)	M(SD)
Feeling about the quality of online discussion tasks & content	4.97 (.78)	4.94 (.92)	4.95 (.88)	4.95 (.86)
Feelings about the quality of online collaborative learning	4.83 (.96)	4.89 (.80)	4.8 (.89)	4.84 (.88)
Attitudes about group structure & membership	3.31 (1.49)	3.78 (1.26)	3.75 (1.21)	3.61 (1.32)
Opinions about assigning a group leader in the discussion form	3.88 (1.63)	3.56 (1.59)	3.8 (1.62)	3.75 (1.61)
Opinions about the value of online discussion & professional development	4.78 (1.01)	5.00 (.59)	4.65 (.88)	4.81 (.83)
Attitudes toward future participation in forums of other online courses	4.50 (1.59)	4.67 (.71)	4.2 (.92)	4.45 (1.07)

Summary and Conclusions

Apparently, consistent differences in online communnication and collaboration can be identified in terms of personality type. Students that tend to be more socially outgoing and engaging, more inclined to agreeableness and seeking of intellectual and/or imaginative experiences seem to be better able to meet the goals of collaborative online discussions, sharing and seeking ideas, comments, questions and concerns. Learning outcomes reflecting cooperation and interactive participation, to some extent, may depend on personality type. Given the previously noted cultural and social roles of communication, the potential effectiveness of online collaborative learning involving discussion and dialog, as opposed to non-interactive, individual responses or monolog, may be limited by personality type. For learning tasks involving interaction and colloboration among group members, the low personality students should not be placed in the same group. As evidenced by this research, placing high and low personality students together in the same group will help accomplish instructional goals related to collaboration and online discussion.

The results of this study suggest grouping strategy effects the interaction, type and pattern of online communication and task engagement. When forming online learning communities for collaborative learning, personality factors could be taken into consideration to promote the potential effectiveness of online communication among groups.

References

Buchanan, T. (2001). *Online implementation of an IPIP Five Factor Personality Inventory* [On-line]. Available: http://www.wmin.ac.uk/~buchant/wwwffi/introduction.html [2002, August 14].

Buchanan, T., Goldberg, L. R., & Johnson, J. A. (1999, November). *WWW personality assessment: Evaluation of an online Five Factor Inventory*. Paper presented at the 1999 Society for Computers in Psychology Conference. Los Angeles, CA.

Buchanan, T., & Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the World-Wide Web. *British Journal of Psychology, 90*, 125-144.

Campbell, W. E., & Smith, K. A. (1997). *New paradigms for college teaching*. Edina, MN: Interaction Books Company.

Comeaux, P. (Ed.) (2002). *Communication and collaboration in the online classroom*. Boston: Anker Publishing Co.

DeFleur, M. L., & Ball-Rokeach, S. (1982). *Theories of mass communication* (4^{th}. ed.). New York: Longman.

Goldberg, L. R. (1999b). *International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality and Other Individual differences* [Online]. Available: http://ipip.ori.org/ipip/ [2002, August 14].

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Books Company.

Johnson, D. W., & Johnson, R. T. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Boston, MA: Allyn and Bacon.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1997). *Active learning: Cooperation in the college* *classroom*. Edina, MN: Interaction Books Company.

Saucier, G., & Goldberg, L. R. (2003). The structure of personality attributes. In M. R. Barrick & A. M. Ryan (Eds.), *Personalityand work: Reconsidering the role of personality in organizations* (pp. 1-29). San Francisco, CA: Jossey-Bass.

Sherry, L. (1996, October). Supporting a networked community of learners. *Tech Trend*, 28-32.

Sherry, L., Billig, S. H., & Tavalin, F. (n.d.). Redefining a "virtual community of learners." Retrieved July 13, 2003, from: http://carbon.cudenver.edu/~lsherry/pubs/TT2000.html

The American Heritage College Dictionary (3rd). (1997). Houghton Mifflin Co. Boston, MA.

Blended Learning: A Different Context For Communities

Frances Clem
San Diego State University

Kirsten Hanson
Oracle Corporation

Abstract

Blended learning involves various delivery methods that may interact in surprising ways. This presentation describes a research study conducted at Oracle Corporation to evaluate how the mix and timing of online and face-to-face interaction affects the formation of a learning community and learner motivation in blended training. The researchers present findings on important effects in blended environments including trust development, learners' perception of content focus, and effects of the corporate culture on the learning community.

Introduction

Many organizations and enterprises are gravitating toward blended learning instead of purely online or Face-to-Face. But it is unknown if the dynamics of blended learning environments work the same or differently compared to the components when used alone. Additionally, blended learning includes media such as telephone conferencing that have not been well researched for their effects on the dynamics of the learning experience.

Our basic research question was: "How does the mix of face-to-face and online interaction work to enhance or reduce the motivation and feeling of community among participants in blended learning situations?" Specifically, we looked at the effects of timing and duration of face-to-face interaction in a blended learning environment. We initially decided to study learning communities in a private sector training environment; we were interested in understanding the impact such communities might have on improving overall business performance.

We studied the Oracle Leader Track, a leadership program that consists of four modules, each containing a two-day, face-to-face workshop and virtual opportunities to enhance and practice key skills. The time commitment for each module is approximately 35-40 hours spread across a five-week period, including the two-day workshop. While designed to be taken in sequence, each module can be taken based on a manager's current needs.

In general, about 65-70 percent of the enrolled participants complete the required activities in each of the five-week modules of the Oracle Leader Track. Given the high expectations of the program, Oracle does not expect one hundred percent of the participants to complete the requirements. However, the firm is interested in the effectiveness of the blended learning design, and the involvement rate is one variable to consider when evaluating the overall return on investment of such a program.

Literature Review

The concept of learning communities is not new, and is informed by earlier thought on the development and characteristics of communities in general (for example, see McMillan, 1986). Learning communities may be seen as being similar to communities of practice (Ahearn & Blunt, 2000, p. 25) and have been defined as "collections of individuals who are bound together by natural will and a set of shared ideas and ideals." (Kowch & Schwier, 1997, p. 3)

The transference of learning communities to electronic communications and virtual reality has added additional layers of complexity to their workings (Palloff & Pratt, 1999; Groff, 1995; Russell, 1999). While these new variations clearly exhibit the characteristics of communities in the larger sense (Haythornthwaite, Kazmer, Robins & Shoemaker, 2000) there is disagreement regarding what a true online community is. For example, Clark (1998) included listservs in his definition, while Kowch & Schwier (1997) specifically excluded them because they do not permit all of what they consider to be the requirements of communities, including negotiation, intimacy, commitment, and engagement. Within online learning communities, researchers have directed their attentions primarily to what works and what does not work (e.g., Palloff & Pratt, 1999; Groff, 1995 and 1996; Merrill, 2000; Ahearn & Blunt, 2000). Palloff and Pratt in particular provide a comprehensive structure for developing and managing online learning communities.

While there has been increasing research and theoretical description of online learning communities (Bonk & Dennen, 2000), there are few comparisons between traditional learning communities and online learning communities. Ellis (2001) compared the two types of communities on such factors as types of communication, participation, discussion, and extension of the learning community beyond the forum. Lucking

(2001) collected quantitative data that indicated that traditional (classroom) learning environments generally produce a higher sense of community than online courses, although he noted some exceptions based on teaching style.

Inquiry into the complex dynamics of online learning communities is also relatively uncommon. Generally, the development of relationships with others in the community is seen as being crucial and observers often comment on the strength and intensity of online relationships (Palloff & Pratt, 1999; Harmon & Jones, 2000; Kowch & Schwier, 1997). The role played by the building of relationships has been examined in detail by Haythornthwaite et al. (2000) as well as by Simich-Dudgeon (1999); the latter used the development of relationships as markers for the growth of the community.

A crucial component in the building of communities is trust (Bonk & Dennen, 2000), at least in part because of its necessity in relationship development (Haythornthwaite et al., 2000). McMillan (1986) noted that emotional safety is needed for a feeling of membership, and Merrill (2000) commented on the need for emotional security in encouraging online learning communities. But how will community dynamics, particularly the development of inter-member trust, change with the advent of "blended" learning environments and their variety of media and interaction types (Zenger & Uehlein, 2001)?

McMillan sets the stage by acknowledging that the components of communities are dynamic and interact in subtle and complicated ways. Moore (1991) suggests that blending media results in a "new form of dialogue." (p. 5) Palloff and Pratt (1999) agree, commenting that the dynamics of interaction can change considerably online. Thus the possibility that blended environments can substantially change the dynamics of learning community formation seems likely. Yet there has been little research to date on how online learning communities are affected by the admixture of blended techniques.

Connected to this question is the role played by face-to-face interaction. Palloff & Pratt (1999) noted that face-to-face meetings can help facilitate the building of an online community, but asserted that it would not change the dynamics of the group. However, Ahearn & Blunt (2000) noted that face-to-face meetings increased communication within the community whenever they took place, and Reid (2002) wrote extensively about the centrality of 'face' in both its presence and its absence in online interactions: "The decoupling of faces and names, and time and place, appears to afford students a perceived degree of anonymity . . . The fact that peers are not able to judge them appeared to be more important than the fact that the lecturer was able to do so. This indicates a perception about social relations—indeed power relations." (p. 61)

It was these variations in social interaction introduced by blended learning, and their effect on learning community development at Oracle, that we set out to investigate.

Method

Our study began with a review of the focus group transcripts from the first three modules. These focus groups were undertaken by Oracle to assess learners' summative opinions about the training after having completed one or more modules. Although few direct references to community existed in these transcripts, they did help to inform the creation of our interview protocol. We then interviewed five past participants who varied in years of experience at Oracle, types of positions, and gender. Upon reflection on our interview findings, we designed the interview protocol for an additional focus group of recent module participants. The focus group questions helped to narrow our focus and deepen our investigation of specific items learned in the individual interviews.

We had lengthy discussions about the major themes discovered in the interviews and focus group. Then, we categorized the participant statements into the theme categories. However, we found overlap in our theme descriptions and further distilled our categories to six major themes after extended discussion. This required us to once again carefully review our transcripts and re-categorize the participant statements into their respective theme category.

This phenomenological approach (Creswell, 1998) seemed to be the most appropriate methodology to employ because the intent was to better understand the meaning of "lived experiences" for several individuals. The interview and focus group participants were chosen because of their experience with the phenomenon of community in a blended learning environment.

Findings

Some themes identified in our interviews indicated similar effects to that noted in the literature, however some differed. In particular, we found affective state changes, such as trust, conflicted with the findings of Pallof and Pratt (1999). This led us to contemplate how effective learning communities in an academic environment might have different characteristics than in a private sector environment.
In order of importance, we have identified the following six theme categories for this study:

Value

This theme focuses on the value the learning community provided for both the learner and the company. How did the learning community benefit the learner/employee? How will it help the company? The amount of learner investment depended on the amount of value received from the community. This in turn impacted the amount of value the company received from the employee's participation in the leadership program. Based on the interview data, it could be posited that a challenging, intense corporate culture – such as Oracle's – benefits from instructional events that develop peer communities.

Much of the value that learners received from the community had to do with the extension of their professional networks and contacts. These new contacts, in turn, helped learners see that their own challenges and issues with their jobs were less unique (and thus, less unsolvable) than they may have previously thought. Interactions within the community helped participants develop a wider perspective of the company and its workings, and added to the range of solutions that participants could draw on.

Comments about value from a few participants include:

"Being less inclined toward groups by nature, I saw a lot of people in the course where I could see a lot of value in creating a tighter network with a lot of people."
"One of the things I got out of it was meeting people from the rest of the company and understanding that they deal with the same kinds of challenges as I do."
"Trust is important to the learning community… the bond… one of the things people experienced that they didn't expect to experience… this is true in both online and classroom… is that they discovered they weren't so isolated… they're so focused on their own little division and they don't realize others are experiencing the same problems or dilemmas."

Trust

We found the development of trust to be related to the type of content being taught (i.e., personal v. instrumental) and also the timing of the face-to-face event (i.e., the week in the five-week program when the meeting took place). To reconcile the apparent conflict between the literature, especially Palloff and Pratt (1999) and the interview results, we conjecture that when people know there is not a classroom experience planned, they may more readily develop trust with their peers in the course; however, we learned they seem to withhold trust when they know there is an upcoming classroom event. We pondered: is trust more closely linked to the task, the environment, or to the expectation of meeting one another?

Comments about trust from a few participants include:

"It's hard to build a good, solid, working relationship without some kind of face-to-face."
"I felt more of a connection [to colleagues] in class, but I didn't feel much of a connection online—it was more difficult to form a connection online. It also appeared to me that other people in the course weren't having as much difficulty in doing that; some of the other people in the course seemed like they were more involved than I was."
"I thought the best way of going through the modules was to have everyone in the classroom setting, the sooner the better… for me, I prefer the first week… I think it really brought people together and formed a real sense of community, people got to know each other, they bonded, and then they were more inclined to participate in the online events going forward… you can put a face with a name… and you didn't want to let that person down. It definitely builds commitment."

Motivation

Content, goal congruence, and delivery methods (face-to-face and online) affected learners' motivation and their involvement in the community. For example, if the content involved more self-discovery rather than specific task orientation, learners were often more engaged in the course. Interestingly, we also realized that learners seemed more likely to use the community for instrumental (task-oriented) reasons rather than experiential (social) reasons. This led us to suspect that job-oriented training environments may call forth different motivational factors than academic learning environments, a fact that instructional designers in organizational training environments may do well to consider.

Comments about motivation from a few participants include:

"The value of the online parts was, after you were in class, having a structure that would make you start applying it for a few weeks until it became a natural part of how you manage."
"The relationship building with the people in the class kind of tailed off over time. Everybody had to go back to their job, and we had less participation following the class. The farther you get away from that 3-day period, the more it's tailing off. People aren't as focused on it anymore."
"The effort is going to come from you first. You can get more help, but if you're not going to put yourself into it, it doesn't matter how many people are there to help you. I think there was a good balance between online, classroom, and coaches."

Perceived Focus

We found participants tended to place the most importance on the delivery method with the "highest touch" element. For example, the face-to-face classroom experience was often seen as "the" instructional event while the online and virtual activities were viewed as support elements to the classroom experience. If further research continues to support this hypothesis, it has significant implications for the placement of content within delivery media in blended environments.

Comments about perceived focus from a few participants include:

"I think we ought to look back on it in two or three years time and I'd be less likely to remember the online activities and more likely to remember the classroom."
"The way I looked at it is the virtual activities reinforced what we got out of the classroom, spread it out a little bit wider, going into some nooks and crevices that we couldn't get into the classroom because we, of course, didn't have any time."
"The quality was definitely improved by having the ILT."
"I know in the classroom I was focused 100% of the time on that, and the outside activity about 50% of the time. I tried to really think about things after class. I didn't really just do it in class and stop doing it once I was out. I thought about things a lot even as I was working on things, driving home, all the other parts of my day."

Environment

Organizational culture and environment have a strong effect on the formation and building of a learning community. For example, a culture like that of Oracle Corporation tends to be very action-oriented and allows minimal time for reflection and deeper collaboration with classmates, which is the primary time when learning occurs. Hence, the cultural norms can have an impact on the level of commitment to a learning community in a blended learning design. These impacts act as constraints on the formation and functioning of learning communities in corporate environments and change the dynamics of learners' expected benefits in exchange for their investment of time in the community.

Comments about environment from a few participants include:

"Some of us in the class wanted to have an online discussion group to continue discussing Smart Goals, but after we left, I don't think anybody ever did. So it didn't become an active community after the class… I think it's a question of the amount of time that people have available for it."
"I think the BYBN case study would have been better face to face than on a conference call… what ended up happening because of the Oracle culture was that we didn't have full participation on the online… and again it's back to commitment. Face to face we would probably have had more individuals participating. People for whatever reason will blow off a conference call rather than a face-to-face commitment. Again, you're talking about a culture shift here."

Personality

It became evident in our findings that, because different personality types have different needs and require different instructional methods, the importance of and investment in the learning community can vary widely from learner to learner. Similar to the impact of environment, this has implications on how participants tend to view the relevance and value of the learning community, and thus the time and effort they are willing to invest in it. This, in turn, may influence instructional designers' determination of how much emphasis to place on community interaction within a blended learning experience, and the level of interaction to require of participants.

Comments about personality from a few participants include:

[Speaker has identified himself as an introvert] "I would tend to learn better individually; certainly I felt the community piece when I was in the class because we were kind of forced together, but online, for me, the community aspect wasn't as strong because my natural tendency to learn individually [meant] that there wasn't a sort of compelling aspect in the community online."
"I think people were more open after the ILT… I really trusted these people with something very personal to myself… also throughout all the modules it was great to see people return… you were hoping those people would come back [to the next module]… you became friends."
"I loved to read the [online bulletin board]. See, I'm a nosy person, so it was like a gossip board for me. People fascinate me."

Conclusions

This was an exploratory study that examined how employees felt about their experience with a learning community. While community formation was not the primary objective of this leadership development program, it was a definite outcome noticed – whether valued or not – by each participant interviewed.

This study of the "community" phenomenon in a blended environment may indicate that not all communities require the same characteristics in order to be effective. The dynamics of learning communities may change with the subject matter, delivery methods, the timing of delivery methods, and the combination in which they are used. This could imply that delivery media may interact with each other and with different domains in ways that are not yet understood. The learner's "internal environment" as well as the learner's "external environment" may also exert a strong influence on how the learner experiences the learning community. For example, such characteristics as personality and motivation level may be equal in importance to the organization's environment (culture) when it comes to determining the success of a community.

The difference between an academic setting and corporate setting may have a more significant impact on a community than originally thought. Most existing research tends to be about communities in an academic setting, and those findings may not be as relevant when creating communities in the private sector. Learners may demand a higher cost-benefit ratio (i.e., the ratio of perceived cost in terms of time and effort compared to perceived benefit in terms of value received) in a job-oriented environment as opposed to a non-working environment. This may be due to the "opportunity cost" associated with carving time for the course out of a day that is already full and demanding.

Finally, the "community" phenomenon is not the only phenomenon at play in this study. None of these employees – as is the case in most corporate training settings – had previously experienced a blended media approach to learning at Oracle. Traditionally, learning, especially with leadership development, has been considered ideal when in a face-to-face setting. Online learning is usually considered to be second-class and less important. Hence, it is probably not surprising that many think virtual learning communities are less effective. So, the initial perceptions and attitudes of learners/employees should be kept in mind when analyzing the findings from this study.

Issues

This limited study allowed for some interesting qualitative data to be gathered, however, the sample size was small. It was also challenging to identify discreet categories in the analysis of the data, hence overlapping, non-exclusive themes (as described in the preceding section) were recognized and should continually be re-evaluated and tested in future studies.

Defining a learning community in the context of Oracle's training was a challenge. The content in each module of the Leader Track was different, creating different levels of community depending on the nature of the content (i.e., assessment of personal values v. functional use of coaching tools). This made it difficult to compare each module's learning community. One could posit that different domains generate different community dynamics and therefore cannot be generalized to all types of learning communities.

Since this research was done in one private sector company, the results are not generalizable to other environments but may possibly be used to inform future research efforts about the impact of blended learning solutions. Finally, researcher bias needs to be considered when reviewing the findings, as one of the researchers is a company employee and has assisted with the design and development of the Oracle Leader program. However, it should be noted that precautions were taken to insure objectivity in the data analysis.

Future Research

Although the results are not generalizable, we hope our approach can be used elsewhere, and that the general conclusions from the study will be useful for informing how communities might be developed in blended learning environments, especially in corporate settings.

Based on our conclusions, it would be interesting to study how different domains or subject matters may interact differently with a blended media mix and how the communities might interact differently. Additional studies on how different personalities and different corporate cultures may encourage or prohibit the development of learning communities might also provide useful suggestions as to the structure of learning designs. Such research could help corporate learning structures and company employees evolve from a more traditional face-to-face community environment to a more networked community that provides support and knowledge exchange beyond the life of the learning event.

References

Ahearn, C., & Blunt, R. (2000). *Creating a virtual learning community.* Paper presented at the Distance Learning 2000: Proceedings of the Annual Conference on Distance Teaching & Learning (16th, Madison, Wisconsin, August 2-4, 2000). (ERIC, ED456235)

Bonk, C. J., & Dennen, V. (2000). *More advances in Web pedagogy: Fostering interaction and the online learning community.* Paper presented at the Distance Learning 2000: Proceedings of the Annual Conference on Distance Teaching & Learning (16th, Madison, Wisconsin, August 2-4, 2000). (ERIC, ED456235)

Clark, C. J. (1998). *Let your Online Learning Community Grow: 3 Design Principles for Growing Successful Email Listservs and Online Forums in Educational Settings*. Retrieved October 11, 2003, from http://www.thehumanizer.com/beta/readings/pop.html

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: SAGE Publications.

Ellis, A. (2001). *Student-Centred Collaborative Learning via Face-to-Face and Asynchronous Online Communication: What's the Difference?* (ERIC, ED467933)

Groff, W. H. (1995). *Creating and Sustaining Learning Communities.* (ERIC, ED389890)

Groff, W. H. (1996). *Creating and Sustaining Learning Communities in the Digital Era.* (ERIC, ED396188)

Haythornthwaite, C., Kazmer, M. M., Robins, J., & Shoemaker, S. (2000). Community Development Among Distance Learners: Temporal and Technological Dimensions. *Journal of Computer-Mediated Communication.* Retrieved October 12, 2003 from http://www.ascusc.org/jcmc/vol6/issue1/haythornthwaite.html

Kowch, E., & Schwier, R. (1997). *Building Learning Communities with Technology*. (ERIC, ED405857)

Lucking, B., Royal, F., & Ireland, M. (2001). *Classroom community in post-secondary classes: An examination of traditional and distance learning environments*. (ERIC, ED 457836)

McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology, 14*, 6-23.

Merrill, H. S. (2000). *Building online learning communities in a graduate program in adult education.* Paper presented at the Distance Learning 2000: Proceedings of the Annual Conference on Distance Teaching & Learning (16th, Madison, Wisconsin, August 2-4, 2000). (ERIC, ED456235)

Moore, M. G. (1991). *Computer conferencing In the context of theory and practice of distance education.* Paper presented at the Applications of Computer Conferencing to Teacher Education and Human Resource Development. Proceedings from an International Symposium on Computer Conferencing (Columbus, Ohio, June 13-15, 1991). (ERIC, ED337705)

Palloff, R. M., & Pratt, K. (1999). *Building Learning Communities in Cyberspace: Effective Strategies for the Online Classroom.* San Francisco: Jossey-Bass.

Reid, C. (2002). *'Facelessness' and its impact on democracy and diversity in virtual communities.* Paper presented at the Diversity/Equity. [SITE 2002 Section]. (ERIC, ED472229)

Simich-Dudgeon, C. (1999). *Interpersonal Involvement Strategies in Online Textual Conversations: A Case Study of a Learning Community.* (ERIC, ED435169)

Zenger, J., & Uehlein, C. (2001). Why blended will win. *T+D, 55*(8), 54-60.

Investigating Strategies for Increasing Student Response Rates to Online Delivered Course Evaluations

Cynthia Conn
Northern Arizona University

John Norris
Northern Arizona University and
University of Maryland

Abstract

Student evaluations of college instruction provide one critical source of information for the improvement of courses, curriculum, and practitioners' pedagogic efforts. Student course evaluations are also often used as one source of data for faculty performance evaluations. Of course, many other sources of information should be tapped via a variety of evaluative methods in order to provide comprehensive and trustworthy evidence for purposes as important as tenure and promotion and course improvement. However, where student evaluations are to be employed in this regard, steps need to be taken to ensure that useful data are gathered and thoughtful student feedback is facilitated.

There is a growing trend towards the use of online-delivered course evaluations for both distance as well as traditional, face to face courses. While online-delivered course evaluations offer many benefits such as lower costs and more timely feedback to instructors, there are also limitations. One of those limitations has been lower response rates. Two studies were conducted to investigate strategies for improving response rates to online-delivered course evaluations. The purpose of these studies was to provide empirical evidence for policies and practices rather than accepting the common perception that response rates for online-delivered course evaluations and surveys will always be low. The findings indicated that a combination of simple and easy to implement strategies were associated with considerably higher student response rates to online-delivered course evaluations than had been the case at Northern Arizona University.

Introduction

Student evaluations of college instruction are one source of potentially useful information on instructional features which contribute to or detract from student learning. Obviously, this is only one method of gathering information for the evaluation and improvement of instruction (Angelo, 1998; Angelo & Cross, 1994; Messick, 1999), but it is a critical one. At Northern Arizona University (NAU) like other institutions of higher education, the use of student course evaluations is also one source of data for faculty performance evaluations. Arizona Board of Regents and NAU Faculty Senate policies require that student feedback be collected and considered during faculty tenure and promotion reviews. Finally, given this requirement, course evaluations can also provide one existing method for collecting student self-assessment data related to learning outcomes, a valuable source of information for instructional improvement.

With the current trends toward electronically mediated instruction, the delivery of course evaluations via a web-based application is mandatory. In addition, many colleges and universities are also considering or implementing course evaluations via an online format for face-to-face courses in order to lower costs, increase timeliness of feedback, and assist with ease of record-keeping and analysis. A key preliminary concern with using online course evaluations for both types of course is whether student response rates will prove dramatically less than has been the case for paper-and-pencil evaluations administered to students in face-to-face classes.

A variety of limitations have been identified in the literature related to low response rates for online-delivered survey instruments. Some of these limitations include technical problems with the tool (Bradley, 1999), difficulty accessing open computers in campus laboratories, and students' technological illiteracy (Handwerk, Carson, & Blackwell, 2000). Others findings have showed that students' multiple e-mail addresses and the frequent changing of e-mail addresses (Bradley, 1999) and slow connections rates (Crawford, Couper, & Lamias, 2001) may also decrease response rates.

Although the challenge of low response rates is clear, a number of benefits can be realized through the use of online delivered course evaluations and survey instruments. These benefits include time and cost savings, quick return of responses, longer and more favorable responses to open-ended questions, and a convenient method for participants who are becoming increasingly technologically literate. Using the Internet to deliver surveys or evaluations eliminates time-consuming and expensive printing processes on the front-end, and it is also an environmentally friendly alternative. In addition, data input of web-based surveys is more efficient and less expensive than inputting responses manually or even scanning paper-based responses (Franceschini, 2000;

Mertler, 2002; Schmidt, 1997; Tse, 1998). Of course, costs can be more expensive if significant time and money is spent building a custom tool. However, if those costs are spread over a number of projects, the initial upfront development may prove less expensive than using a commercial product.

Given these potential benefits, NAU's Office of Academic Assessment (OAA) decided to investigate strategies for improving response rates to online-delivered course evaluations. The purpose of this study was to provide empirical evidence for policies and practices, rather than accepting the common perception that response rates for online-delivered course evaluations and surveys will always be low. OAA undertook two studies to investigate this topic. The first study was intended to determine how low response rates actually were for course evaluations delivered through a proprietary web-based tool and what factors might increase responsiveness. The second study explored the effectiveness of a combination of very simple strategies for increasing response rates. An overview of findings for study 1 and the methodology and findings for study 2 are described in the following report (full details can be found in Norris & Conn, 2003).

Study 1: Overview of Findings

Study 1 sought to establish a baseline of student response rates to online course evaluations at NAU by analyzing data from an existing online course evaluation system, Evalajack, which has been in use since the Spring semester of 2000. Evalajack is a proprietary web-based tool that was developed by NAU Information Technology Services (ITS) to provide an immediate, if not comprehensive, solution to the burgeoning demands for online evaluations in conjunction with rapidly increasing web-based course offerings. Between Spring 2000 and Fall 2002, Evalajack forms were made available for over 1000 web-based course sections at NAU. In this study, we first analyzed existing response rates data for these two years of online course evaluations; we then followed up with a questionnaire (see Appendix A) to a sub-set of online instructors in order to identify variables which might be associated with higher or lower student response rates.

Findings

Analysis of existing data for Evalajack online evaluations indicated generally low student response rates. Table 1 shows descriptive statistics for response rates, with course section taken as the unit of analysis. Based on data for 1108 course sections, an average of only 31% of students per course responded to end-of-semester evaluations, although a standard deviation of 28% also indicated considerable variability from course to course. Furthermore, on average, 11% more students responded to evaluations in graduate courses than did in undergraduate courses, while variability was equally high in each of these sub-sets of data.

Table 1. *Response rates for Evalajack, Spring 2000 – Fall 2002*

Course Sections	Statistic	Response Rates
	Mean	31%
Combined	**S**	28%
(N = 1108)	Min	0%
	Max	100%
	Mean	36%
Graduate	**S**	28%
(N = 625)	Min	0%
	Max	100%
	Mean	25%
Undergraduate	**S**	27%
(N = 483)	Min	0%
	Max	100%

These findings should not be over-interpreted. Given the anonymity and multi-year nature of the data, it was impossible to confirm with individual instructors whether Evalajack forms had been made available to students in every course section, or whether other problems were encountered in accessing and completing the forms. In addition, it may well have been the case with a subset of the data that alternative evaluations (e.g., instructor-created forms) were used in lieu of, or in addition to, the Evalajack forms for any number of reasons. Nevertheless, given the fact that evaluations were created and made available online, and that instructors were requested to announce the evaluations to their students, it is reasonable to interpret the current findings as evidence that student response rates to Evalajack online course evaluations have been, on average, relatively low. However, due to unknown factors, a sub-set of courses/instructors did manage to encourage reasonable to good levels of responsiveness from students, and this finding was not largely attributable to class size or undergraduate/graduate classification of the course.

Results from the Fall 2002 instructor questionnaire helped to identify factors potentially associated with differences in response rates to Evalajack online evaluations. Given the low number of instructors queried (85) and responding (50), definitive associations could not be drawn between specific variables or combinations of variables and increases in response rates; multivariate inferential statistical assumptions were simply not met by the low number of participants in this study (cf. Tabachnik & Fiddell, 1996). However, patterns in their answers to the questionnaire did emerge, and these tentative relationships with increasing response rates were used to inform further investigations.

These patterns indicated several potential strategies for increasing student response rates to online evaluations. The timing and location of evaluation announcements was apparently associated with response rates, with advantages for announcements between two and three weeks prior to the end of the semester and the posting of initial announcements via online discussion boards and course management web sites. In addition, posting of evaluation announcements in multiple locations and/or the use of reminder notices seemed to increase student responses. Instructors' open-ended feedback also served to further contextualize these possibilities. First, several instructors commented on the incompatibility of Evalajack forms with certain web browsers, and they stated that a stable and accessible evaluation platform was a prerequisite for enabling students to respond. Second, a number of instructors noted that they were worried about "bombarding" students with too many reminders to complete an evaluation, as well as with too many evaluations (e.g., from Statewide programs as well as from their college), and they were concerned that students might react negatively within evaluations as a result. Finally, several instructors commented that low response rates to online course evaluations posed a serious problem for Tenure & Promotion review, and they added that evaluations should ask questions directly related to the course and comparable to evaluations of other courses within the college/school, in order to be meaningful to student respondents as well as to faculty and administrative users. All of the findings above, and in particular the insightful open-ended comments by instructors, informed the development of strategies investigated within the subsequent study.

Study 2: Method and Findings

Study 2 investigated the effectiveness of a set of combined strategies intended to increase student response rates to online course evaluations. Based on factors observed in study 1 to potentially influence student responsiveness, simple recommendations were compiled for instructors to use in announcing, discussing, and reminding students to complete the online course evaluations. Several constraints influenced the strategies selected. First, given the potential positive or negative effect on instructor evaluations, penalty/reward strategies were not recommended. Second, in order to minimize potential technological influences on response rates, an alternative and reliable online evaluation system was utilized in lieu of Evalajack. Third, in response to instructors' concerns with over-burdening and annoying students, it was decided that only a minimal set of strategies would be operationalized based on (a) ease of implementation across widely varying individual course contexts (e.g., wherein different web-based course management tools were used) and (b) likelihood of increasing response rates. Finally, because large numbers of faculty/instructors were not available for participation in the study, it was decided that individual strategies could not feasibly be investigated in isolation. In addition, because the course evaluation process and results would be taken seriously and used for various purposes by students, instructors, and administrators, ethical concerns dictated that all courses should be treated equally and that the study intervention should attempt to maximize student response rates. Accordingly, only a single set of combined strategies was operationalized in all courses in the current study.

Method

Two distinct sub-sets of courses/instructors participated in the study. First, voluntary participation in the study was solicited from NAU faculty and instructors teaching fully web-based courses during the Spring or Summer 2003 semesters. Participants were solicited from those colleges/schools/departments where Evalajack was the only online evaluation tool in operational use at the time (several schools had implemented other in-house evaluation tools in order to meet unique requirements). Twenty-one instructors and faculty, teaching a total of 39 distinct course sections (22 graduate, 17 undergraduate), volunteered from a variety of disciplines in four colleges/schools: Arts and Sciences, Business Administration, Education, and Social and Behavioral Sciences. The number of students enrolled in these online courses ranged from 3 to 49, with an average of 24 students per section. Courses were all taught entirely online during 5-week, 10-week, or 16-week semesters. Participants agreed to the following conditions of the study: (a) the OAA would design and host course evaluations; (b) evaluations would include all items from the required college/school forms plus any additional items about the course or student learning outcomes that the instructor requested; (c) participants would follow a set of basic strategies, provided by the OAA, for increasing student response rates; and (d) anonymous results would be reported to the instructor and other required parties (e.g., department chairs) within one week following the grades due date for the semester.

In addition to these web-based courses, instructors for University Colloquium 101, a required

Freshman-level course that is taught in small face-to-face sections, were invited to participate in the response rates study during the Spring 2003 semester. Given the very large number of UC101 sections offered each semester at NAU, the administration and the OAA had decided to investigate the possibility of migrating the end-of-semester course evaluations to an online rather than paper-and-pencil administration format, in order to facilitate gathering, analyzing, and reporting student responses. Of the 25 sections offered during Spring 2003, instructors of 21 sections agreed to participate in the study; enrollment in these sections ranged from 19 to 30 students, with an average of 23 per section. For comparison purposes, additional response rates data were collected for paper-and-pencil evaluations in the UC101 course for both Fall 2002 and Spring 2003 semesters, and response rates for pilot online evaluations from the Fall 2002 semester (when no strategies for increasing responsiveness had been implemented) were also included.

In order to ensure that technology problems (e.g., incompatibility between evaluation software and certain web browsers) would not influence response rates, a stable web-based evaluation system was identified in the commercial survey-design and -hosting software, SurveyMonkey (http://www.surveymonkey.com). The OAA had pilot-tested several online course evaluations using this software during Fall 2002, and instructors/students had reported very few difficulties in accessing or completing the evaluations on the SurveyMonkey website. Therefore, end-of-semester evaluations for all participating courses were designed and hosted on this stable platform.

After evaluation forms had been designed and uploaded, a unique URL was provided to each instructor for each web-based course section. Participants were then asked to implement the following strategies (exact instructions are provided in Appendix B). First, the evaluation URL was to be posted, and clearly identified, in a prominent location on the course management home page two weeks prior to the last day of class for the 16-week semester (times were adjusted for the shorter summer semesters). Second, students were to be notified simultaneously of the availability of the course evaluation and the URL for completing it via postings in all locations appropriate to a given class (including e-mail, the online course syllabus or course calendar, and online course discussion boards). Participants were also asked to include in these announcements a brief statement regarding the value of completing course evaluations, as well as instructions to students for doing so and a date by which evaluations should be completed. Finally, participants were requested to send to their students, one week prior to the final day of classes, an e-mail message reminding students to complete the evaluation and including the URL and 'complete-by' date. A single e-mail reminder strategy was selected over other possible reminders, due to the fact that all instructors utilized e-mail for regular communication with their online students but not necessarily other web-based tools (e.g., discussion boards). In addition, given the proximity of the reminder to the initial announcement (one week following), multiple reminders were not recommended in order to avoid potential negative reactions from students.

The implementation of strategies in the UC101 sections differed in several ways due to the face-to-face course format. Instructors were provided with an evaluation instruction handout to distribute to students between two and three weeks prior to the last day of classes. These instructions informed students about the value of completing course evaluations, provided them with the online evaluation URL and instructions for accessing it outside of class, and stated the date by which forms should be completed. Instructors were asked to announce the online evaluation in class and to spend a few minutes discussing the instructions. One week prior to the last day of classes, the OAA sent an e-mail reminder message to all students enrolled in UC101; this message requested that students complete the evaluation if they had not already done so, and it provided a 'clickable' URL link to the form for each course section. After the semester ended, a questionnaire was sent to all instructors to determine if the outlined strategies were followed. The responses showed that the strategies were adhered to for the most part and there were very few deviations.

Findings

Overall, considerably higher average student response rates to online evaluations were found for courses implementing the simple strategies outlined above than had been observed for Evalajack online evaluations over the previous two years. Table 2 shows response rate data for the sub-set of web-based courses that participated in the study, both combined and broken down according to graduate/undergraduate classification. In order to provide an accurate indication of the extent to which strategies may have influenced student responses, existing response rate data from previous administrations of Evalajack were also analyzed for the same sub-set of classes on their own (i.e., rather than comparing the overall Evalajack average response rate of 31% with this sub-set). For these web-based classes, average response rates with the strategies were 32% higher than without, up from 42% to 74%. Likewise, average response rates for graduate courses were 29% higher with strategies and for undergraduate courses, 35% higher with strategies. Minimum and maximum response rates for individual sections were also consistently higher under the strategies condition.

Table 2. *Response rate descriptive statistics for evaluations in participating web-based courses*

Courses	Statistic	No strategies (Evalajack)[a]	Strategies
Combined	N	37	39
	Mean	42%	74%
	S	22%	18%
	Min	6%	33%
	Max	81%	100%
Graduate	N	21	22
	Mean	44%	73%
	S	22%	18%
	Min	6%	38%
	Max	77%	96%
Undergraduate	N	16	17
	Mean	39%	74%
	S	23%	18%
	Min	9%	33%
	Max	81%	100%

[a]Note. "No strategies" statistics were calculated based on response rate data available from all previous administrations of Evalajack online evaluations; data were unavailable for two of the participating courses which had not previously administered online evaluations.

Similar patterns were observed for response rates in the UC101 face-to-face courses. Table 3 shows response rate analyses for three groups of UC101 sections: (a) the Fall 02/Spring 03 sections that administered paper-and-pencil evaluations in class (Paper 02-03); (b) the Fall 02 sections that administered online evaluations prior to the strategies study (E 02); and (c) the Spring 03 sections that administered online evaluations and implemented the recommended strategies (E 03). Once again, the average student response rate for sections that implemented the strategies was found to be much higher than for the pre-strategies sections (up from 34% to 67%). Note also that the minimum response rate for any section using the strategies was observed to be 41% (compared with 0% for the pre-strategies sections), and the maximum was 96%. While the average response rate for paper-and-pencil sections was found to be 83% (16% higher than the online evaluations with strategies), this difference had been anticipated, given the fact that none of the UC101 course sections was taught in a web-based format. In other words, an average of 67% of the freshmen students in this face-to-face course responded to the request that they complete an online evaluation of the course at the end of the semester outside of class time, a rather remarkable finding. Of additional interest is the fact that 100% response rates were not found for the UC101 sections that elected to administer paper-and-pencil evaluations in class; indeed, a number of the online + strategies sections received higher response rates than did several of the paper-and-pencil sections. Clearly, absenteeism in the paper-and-pencil sections played a role in this observation. Finally, it should be noted that three additional UC101 sections that had elected to administer paper-and-pencil evaluations failed to do so, for a response rate of zero (data for these sections were not included in the calculation of average response rates).

Table 3. *Response rate descriptive statistics for UC101 evaluations*

UC101 sections	Statistic	Response Rate
Paper 02-03 (*N* = 11) [a]	Mean	83%
	S	13%
	Min	63%
	Max	100%
E 02, no strategies (*N* = 48)	Mean	34%
	S	18%
	Min	0%
	Max	75%
E 03, strategies (*N* = 21)	Mean	67%
	S	17%
	Min	41%
	Max	96%

[a]Note. Three additional sections elected to administer paper evaluations but did not return any completed forms (i.e., 0% response rates). Data from these sections are not included in the current response rate analyses.

It should be obvious that the research reported here was not designed to inform causal interpretations about the relative effectiveness of individual or combined intervention strategies at increasing student response rates; such interpretations would only be warranted through the operationalization of individual strategies in a large number of carefully sampled courses and the control of other potentially intervening variables. In addition, a variety of unexamined factors might have influenced the different response rates observed within the courses in the current study, including: (a) overall consistent increases in students' online literacy and familiarity from semester to semester; (b) the variable interactiveness and engagement of students and instructors in different online courses; (c) differential perceived worth of the course evaluation process among students; (d) interaction effects between user-side technology and online evaluation technology; and many others.

However, despite these obvious limitations in the current study, findings clearly indicated that students responded with relatively high average frequency to online course evaluations that were administered via a stable platform and in conjunction with a very simple set of strategies to encourage their responsiveness. It is also clear that response rates for online courses which employed these strategies were substantially higher than for other online courses that did not, as well as for the same courses prior to the use of such strategies, and this pattern was reproduced for an online-delivered course evaluation used in a set of face-to-face classes that had no online instruction whatsoever. Perhaps most encouraging was the observation that, on average, 74% of the students enrolled in web-based courses responded to the course evaluation when simple strategies were employed, a response rate that was only 9% less than that observed for paper-and-pencil evaluations administered in the face-to-face UC101 class. Furthermore, it should not be overlooked that average response rates for the UC101 online evaluations without strategies (34%) were very similar to average response rates found for Evalajack overall (31%), this despite the fact that a stable online platform was used to host the UC101 evaluations. This finding lent considerable support to the interpretation that strategies were the primary contributor to the observed differences in response rates (i.e., rather than the simple reduction of potential technological problems).

Conclusion

Student evaluations of college instruction provide one critical source of information for the improvement of courses, curriculum, and practitioners' pedagogic efforts, and their use as a component of faculty review is a well-established tradition in higher education not likely to disappear any time soon. Of course, many other sources of information should be tapped via a variety of evaluative methods in order to provide comprehensive and trustworthy evidence for purposes as important as tenure and promotion, course improvement, and others. However, where student evaluations are to be employed in this regard, steps need to be taken to ensure that useful data are gathered and thoughtful student feedback is facilitated.

In the studies reported above, we found that a combination of simple and easy to implement strategies were associated with considerably higher student response rates to online-delivered course evaluations than had been the case for prior online evaluations at NAU. Average positive differences of 32% were found for online courses implementing these strategies, and the resulting average response rate of 74% approximated closely the rate found for paper-and-pencil evaluations administered in multiple face-to-face sections of a freshman-level class (83%). Similar positive differences were also found for the use of these strategies in conjunction with online-delivered evaluations in multiple sections of a face-to-face course.

Thus, while the mere creation and posting of an online course evaluation is probably insufficient for ensuring that students will respond to it—as clearly demonstrated in the very low average response rates found for existing online evaluation data at NAU prior to this study—it is not necessarily the case that this situation cannot be rather easily ameliorated. An obvious prerequisite in this respect is the use of a stable and accessible online platform for designing and hosting course evaluations, such that students encounter minimal technological impediments to the evaluation process. In addition, it would seem that a few simple efforts may go a long way towards successfully encouraging considerable proportions of students to provide their critical feedback, including: (a) explicitly announcing the availability and location of the evaluation within a few weeks of the end of the course; (b) explaining the value of the course evaluation process and student feedback; and (c) reminding students to complete the evaluation. It may be, as was found in the current research, that these minimal efforts (similar, by the way, to the most fundamental actions we take as instructors in assigning homework or administering an online quiz) will prove sufficient for gathering adequate levels of student response data via online-deliverd course evaluations, without the addition of dubious coercion or reward methods (likely to have systematic negative or positive effects on course evaluation outcomes).

In the end, it is our hope that paying due attention to such basic concerns will result in improvements in the increasing use of online-delivered course evaluations. Just as effective online instruction requires more than simply posting a syllabus and course readings to the web, so too will effective online course evaluation require something more than simply posting the evaluation URL to a course management web site. Finally, once we are able to move beyond concerns with student response rates, the real work of developing useful course evaluation instruments which will inform reliable and valid interpretations about instruction can begin.

Acknowledgements

We would like to thank the faculty members and instructors who participated in the research reported here, as well as the many students who responded to their online course evaluations. Thanks also go to Brian Sawert at NAU Information Technology Services for his assistance in querying the Evalajack database. Finally, we appreciate the support, encouragement, and feedback offered by Paul Rowland during early phases of this work.

References

Angelo, T. A. (Ed.). (1998). *Classroom assessment and research: An update on uses, approaches, and research findings*. San Francisco: Jossey-Bass.

Angelo, T. A., & Cross, P. K. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.

Bradley, N. (1999). Sampling for internet surveys: An examination of respondent selection for internet research. *Journal of the Market Research Society, 41*(4), 387-395.

Crawford, S.D., Couper, M.P., & Lamias, M.J. (2001). Web surveys: Perception of burden. *Social Science Computer Review, 19*, 146-162.

Franceschini, L.A. (2000). *Navigating electronic survey methods: Three pilot studies*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association. Bowling Green, KY.

Handwerk, P.G., Carson, C., & Blackwell, K.M. (2000). *On-line vs. paper-and-pencil surveying of students: A case study*. Paper presented at the Annual Forum of the Association for Institutional Research. Cinncinnati, OH.

Mertler, C.A. (2002). *Web-based surveys: Guiding lessons for their use*. Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA.

Messick, S. J. (Ed.). (1999). *Assessment in higher education: Issues of access, quality, student development and public policies: A festschrift in honor of Warren W. Willingham* . Mahwah: Lawerence-Erlbaum Associates.

Norris, J. M., & Conn, C. (2003). Investigating strategies for increasing student response rates to online-delivered course evaluations. Manuscript submitted for publication. Draft copy retrievable online at http://www4.nau.edu/assessment/oaainfo/publications/publications.htm

Schmidt, W.C. (1997). World wide web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, and Computers, 29*(2), 274-279.

Tabachnick, B.G., & Fidell, L.S. (1996). *Using multivariate statistics* (3rd ed.). Boston, MA: Allyn and Bacon.

Tse, A. (1998). Comparing the response rate, response speed and response quality of two methods of sending questionnaires: E-mail vs. mail. *Journal of the Market Research Society, 40*(4), 353-361.

Appendix A

Questionnaire: Online Course Evaluation Response Rates

Please provide us with the following information about your end-of-semester course evaluations for Fall, 2002 online courses.

Please list the course names and section numbers for those courses to which your answers on the following questions apply:

1. When was the end-of-semester course evaluation first announced to students?
   a. 1 week before the last day of the semester
   b. 2 weeks before the last day of the semester
   c. 3 weeks before the last day of the semester
   d. More than 3 weeks before the last day of the semester
   e. Other (please specify):

2. How was the end-of-semester course evaluation first announced to students? (Check all that apply)
   a. In class verbal announcement (hybrid class)
   b. Paper handout (hybrid class)
   c. E-mail announcement
   d. Posted to class web site (online or hybrid class)
   e. Posted to class assignments (online or hybrid class)
   f. Posted to class discussion board (online or hybrid class)
   g. Other (please specify)

3. Did you discuss the value of end-of-semester course evaluations with your students?

   If yes please explain when and how:

4. Did you use any reminder strategies to encourage students to respond to the evaluation? (Check all that apply)
   a. In class verbal reminder (hybrid class)
   b. Paper handout (hybrid class)
   c. E-mail reminder
   d. Posted reminder to class web site (online or hybrid class)
   e. Posted to class discussion board (online or hybrid class)
   f. No reminder strategy used
   g. Other (please specify)

5. Did you offer students credit (participation points, extra points, etc.) for completing the evaluation?

   If yes please explain what kind of credit was offered:

6. Did you require students to complete the evaluation (e.g., in order to access the last assignment)?

   If yes please describe the requirement:

7. Did students complete a mid-semester evaluation of your course?

8. Please comment on any additional strategies used for encouraging student responses to online course evaluations and please add any other insights into this issue.

Appendix B

Strategies for increasing response rates to online course evaluations

1. Posting course evaluation link: For WebCT courses, post the link (url) for the course evaluation directly to the course home page, with the heading "End-of-semester Student Evaluation of Course". Make the link available at least 2 weeks prior to the last day of class. Likewise, for courses offered via other course management systems, post the link for the course evaluation in a prominent location on the course home page. DO NOT post the link to the old Evalajack evaluation. (We will also inform Statewide/Distributed Learning not to post the Evalajack evaluation for your course.)

2. Announcement of course evaluation availability: Simultaneous with posting the evaluation link, announce to students the date of initial availability of the course evaluation in each of the following locations, where applicable to your course: (a) syllabus; (b) discussion board; (c) e-mail message; (d) course calendar.

Include a statement of value for the course evaluation process and instructions for completion in the initial announcement. Make sure to also include a "complete by" date in the initial announcement. [NOTE: Course evaluations should always be completed prior to the final day of the course (before grades are turned in).] The following example may be used and adapted/personalized for announcement purposes:

"Course evaluations are very important at NAU. They enable students to provide instructors and the university with feedback about the effectiveness, quality, and value of courses. Your feedback on this course evaluation is essential for revising and improving this course and how it is taught. These improvements will be of direct benefit to future students. Improvements will also benefit you, by raising the overall quality of courses at NAU and increasing the reputation and value of an NAU degree. Finally, I greatly value your feedback, and I appreciate your attention to the evaluation of this course.

> By clicking on the link below, you will access the anonymous end-of-semester evaluation for this course. Please contact d-oaa@jan.ucc.nau.edu (the online evaluation administrator) if you run into any problems with completing the evaluation online.
>
> Please complete the course evaluation by [PROVIDE THE DATE FOR THE LAST DAY OF CLASSES]"

3. Reminder to complete course evaluation: One week prior to the last day of class, send an e-mail reminder to students to complete the course evaluation if they have not already done so. Include the evaluation url link in the e-mail message, and remind students of the "complete by" date.

A Taste of Their Own Medicine: Use of an Electronic Performance Support System (EPSS) by Performance Technologists

Abbas Darabi
Melissa C. Mackal
David W. Nelson
Florida State University

Abstract

This study describes the use of a Web-based Electronic Performance Support System (EPSS), named ePlan™, by students in a graduate course on Performance Systems Analysis(PSA). ePlan™ assists performance consultants in analyzing performance issues, performance gaps, and causes, and provides guidelines for recommending performance solutions. The paper reports ePlan™ utility, its impact as a support tool, and the users' reactions to the EPSS including changes in their self-efficacy to perform the PSA analysis.

Introduction

Complex cognitive skills are defined as skills with a number of components or "constituent skills," the majority of which are performed cognitively (van Merrienboer,1997). Performance analysis refers to a set of skills containing numerous cognitive tasks and therefore can be considered a complex cognitive skill. Because of its complexity, learning to conduct a successful performance analysis requires students to execute goal-oriented behavior, diligent and persistent performance, and continual self-assessment. The present study examines how an EPSS for performance analysts, supported students to develop competence and how it complemented the course activities by promoting self-regulated learning.

Performance analysis is an application of human performance technology (HPT), which the International Society for Performance Improvement (ISPI) defined as "the systematic and systemic identification and removal of barriers to individual and organizational performance" (ISPI, n.d.). For successful, application of HPT, the performance analyst should rigorously adhere to a data-driven methodology, yet flexibly respond to unanticipated cues that arise during the analysis. Analysts' abilities to obtain acceptance for implementation of their recommended solutions are heavily dependent upon their competencies as demonstrated to the clients. In her discussion of required competencies for performance consultants, Hale (1998) includes the following: Identifying the knowledge and skill requirement for the job, defining and describing clear project output, tracking and describing behaviors and their effect, business understanding, project management, time management, systems thinking, communication, group management and teamwork, data reduction, information search, and intellectual versatility. Thus, learning to become a successful performance analyst or consultant is a goal-based behavior in which the precepts of HPT theory, the use of tools, and the knowledge of results from practice are synthesized into an understanding of how the performance analysis should be done.

In a performance analysis course, the initial development of this understanding by students requires a mix of instruction in theory, practice at the use of strategies and tools, and real-world experience. The learning goals for students are to acquire competence and confidence in using HPT strategies and tools in a gradual, cyclical process of self-regulated learning. The following sections of the paper present the theoretical basis of this process and a description of the learning activities of the course.

A Theory of Self-Regulated Learning and Self-Efficacy

The theory of self-regulated learning from the social cognitive perspective arose from Bandura's social cognitive theory (Bandura, 1991), which identified three self-regulatory processes: self-observation, self-judgment, and self-reaction. Self-observation involves learners' perceptions of their own knowledge, skills, strategies, or performances. Self-judgment involves learners' comparisons of their current standings with respect to their personal goals or standards, and their decisions to strive to meet or exceed them. Self-reaction involves their actions to meet or exceed their goals, which might include increased effort or adopting different strategies. As self-regulated learners accomplish their goals, their self-efficacies about the learning tasks – their beliefs that they can succeed in the required tasks, the accomplishment of their personal standards, and the use of strategies and tools – continually increase (Bandura, 1991).

Self-regulated learners generally are flexible in their applications of strategies and attribute their successes or failures to their own efforts rather than to external influences. Having observed that their efforts and strategies have resulted in the accomplishment of their personal standards, they raise their standards in a new cycle that begins with self-reaction. Alternatively, having observed that their efforts did not result in accomplishment of their goals, they alter their strategies or increase their efforts (Zimmerman, 1998).

The Performance Analysis Course

The primary author has repeatedly instructed a graduate level course titled Performance Systems Analysis (PSA) at Florida State University. The main task for students in this course was to conduct a performance analysis project on a real-world performance problem as a member of a team. The course prepared students who had learned instructional design skills to embrace the changes in their field toward performance technology, and to become performance analysts. In this course offering (Spring 2003), the instructor included a requirement to use an EPSS specifically developed for performance analysis. A detailed description of this support system follows in the next section of this paper.

The course guided students to analyze performance systems, identify workforce performance problems and their causes, and recommend appropriate performance interventions. The complexity of the tasks gives good reason to apply a situated learning approach (Brown, Collins, & Duguid, 1989) to the course, during which students were able to coordinate and integrate the constituent skills of performance analysis. This approach made the practical experiences of the course critically relevant to students' future professional performances.

In addition to the foundation of instructional systems literature with which students were already familiar (e.g., Gagné, 1985; Gagné, Briggs, & Wager, 1992; Gagné & Medsker, 1996), students were introduced to supportive literature on HPT and human resource development (e.g., Dean, 1999; Gilbert, 1996; Harless, 1975; Mager and Pipe, 1997; Robinson and Robinson, 1995; Rosenberg, 1990; Rossett, 1999; Rummler, 1999), and literature on management design and systems thinking (Ackoff, 1999; Gharajedaghi, 1999). The course introduced theories, research, and methodologies of HPT.

In accord with the principles of situated learning (Brown, Collins, & Duguid, 1989), the twelve students were assigned to three teams that studied real performance problems. The students were introduced to three clients who requested assistance with solving performance problems in their organizations. The instructor grouped students with complementary skills to address the designated performance problems. Each team was assigned to a project, considering the individuals' preferences for the organizations and the performance issues.

Students were also provided access to an EPSS. Every student was required to use this EPSS and individually enter the information they collected from their clients. They used the methodologies described in class to collect information and to conduct their analyses. They employed quantitative and qualitative methods of data collection such as interviews, observations, focus groups, document analyses, and surveys. The data collected by these methods were entered into the EPSS following the scheduled phases of the project. In these phases students described the desired and actual states of the organizational performance, conducted gap and cause analyses, and developed their recommended solutions. Students also presented their progress and findings to their clients in formal meetings.

EPSS for Performance Analysis

ePlan™ is a web-based Electronic Performance Support System designed and developed by the Learning Systems Institute of Florida State University for the United States Navy. ePlan™ assists performance consultants in the analysis of performance problems that have been identified by clients. ePlan™ draws on a synthesis of several HPT models to facilitate the following analysis activities: identify performance problems, analyze workforce performance, select specific organizational results to be examined, select data collection methods, weigh and prioritize performance gaps to be closed, identify potential causes of performance problems, and identify solutions to close performance gaps. Throughout these activities ePlan™ generates nine reports to be used by analysts and clients.

ePlan™ provides numerous tools to aid the novice performance analyst to investigate causes of, and recommend solutions for, performance problems. It programmatically structures the analysis activities in three phases: Define, Analyze, and Select. It updates the existing information used in activities and reports throughout the use of ePlan™ as analysts enter new data. ePlan™ also allows users access to reports from multiple projects and facilitates a threaded discussion for users to communicate across projects, thus promoting a community of practice.

The support system does not alleviate the responsibility of analysts assigned to a project to conduct the analysis. Instead, it aids the consultants to gather data, produce reports, explain the meaning of the data to clients, and negotiate priorities of gaps, causes, and solution alternatives.

Method

The design of this investigation was a single case study concentrating on how the components of the PSA course enhanced students' competence acquisition and self-regulated learning. Students started the course with an introductory discussion of the HPT concepts and theoretical basis. Once students were assigned to teams and introduced to their prospective clients, all three teams were presented with an overview of ePlan™, its contents and its role in the students' intervention during the span of their team projects. During the course of the semester, teams would individually meet with teaching assistants for a total of three sessions. Each session was to review and complete the three individual sections of ePlan^{TM}: the Define Phase, the Analyze Phase and the Select phase. During the class instruction, teams learned about the tasks involved in each of these phases and started collecting data required for their projects. It was the teams' responsibility to contact the teaching assistants and begin ePlan^{TM} activities. Upon client contact, each team was instructed to review the data they had collected from their analysis of the performance issues, confirm their readiness and schedule a session appointment. Questions and suggestions were provided to assist the students for first session preparation. These recommendations specifically related to the components of each phase and thus differed in nature for each session.

At the beginning of each session, teams were presented with an overview of the session's schedule. Although the students worked with clients in groups and conducted analyses of workforce environments in teams, they were required to individually log into ePlan^{TM} and enter their team data. This flexibility provided students the opportunity to experience the ePlan^{TM} individually while sharing the experience of collecting, analyzing, and interpreting the data as teams. They were encouraged to discuss the issues of their assignments with one another during the use of ePlan^{TM}.

At various points of their work with ePlan^{TM} , students printed out performance reports and handed them to the teaching assistants. This procedure assured that each student participated in all aspects of ePlan^{TM} phases and met its requirements. At the conclusion of the last session, students completed data collection instruments and participated in focus group discussions concerning their experiences with ePlan^{TM}.

Measures

Qualitative and quantitative data collection methods were employed at different points of the students' experiences with the course and ePlan^{TM}. Following is a description of these methods and instruments:

Self Efficacy Measure

Bandura and Schunk (1981) discussed the conceptual framework and application of self-efficacy measure in a study of mathematical efficacy. Based on their application of this measure, we developed a five-item measure of students' self-efficacy of their perceived competence at conducting performance analysis and its components. These items specifically addressed the phases of performance analysis and the inclusion and application of ePlan. Each item presented a scale of 10 to 100 to represent values ranging from "Not sure" to "Very Sure". Items included questions such as, "On scale from 10 to 100, how sure are you that you will be able to successfully conduct a Performance Systems Analysis after completing this course?"

The pretest instrument was administered after the overview of ePlan^{TM} and before students began their ePlan^{TM} projects. The purpose of ePlan^{TM} overview was to give students a preparatory exercise which, alongside the introductory class discussions of conducting a PSA, acquainted them with the subject. The posttest instrument was given to students immediately after completion of their ePlan^{TM} projects.

Student Feedback

This questionnaire was designed to collect information on the usefulness and quality of ePlan^{TM}. Two types of questions were used to measure students' reactions to the quality and technical features of ePlan^{TM}. Questions specifically focused on whether ePlan^{TM} was helpful in formulating their team approach, informative in their preparation, and instructive in guiding them through the phases of PSA projects. One series of questions, constructed on a six-point Likert scale, corresponded to the phases of PSA. A set of open-ended questions provided students the opportunity to describe their personal experiences with using ePlan^{TM} and offer suggestions for its improvement.

Focus Group Sessions

At the completion of the students' ePlan^{TM} projects, each team of students participated in a focus group discussion of their experiences with ePlan^{TM}. These discussions mainly derived from students' previous reactions to the use of ePlan^{TM} throughout the semester. The format of the sessions was casual and the meetings were scheduled flexibly to assure the participation of all team members. Discussions included themes of personal experiences, difficulties using the tool, revisions of the tool, its alignment with the course design, and the relevance of the various features of ePlan^{TM} to real world projects.

The focus group discussions were recorded and later transcribed for analysis purposes. This transcription provided a set of qualitative data that complemented information collected with other instruments.

Data Analysis and Results

The quantitative data collected by self-efficacy instruments and the feedback questionnaire were prepared for analysis using Statistical Package for Social Sciences (SPSS). A simple descriptive analysis was performed on students' reactions data. A paired samples *t* - test was used to analyze the pre and post ePlan^{TM} differences in self-efficacy.

The qualitative data collected through tape recording the focus group discussions were transcribed and merged with the responses to the open ended questions in the feedback survey. They were combined into a list of items describing students' experiences with the use of ePlan^{TM}. The authors first individually reviewed the students' statements and interpreted the fundamental nature of each. They then identified the features of ePlan^{TM} to which the statements referred. Without exception, all the students' negative comments referred to one or more technical aspects of ePlan^{TM} and all of their positive comments referred to the use of the software and the support it provided for the students projects. At the end of this iterative refinement and abstraction process, students' statements were categorized into more general themes which characterized the use of ePlan^{TM} in this PSA course.

We present the results of our analysis in terms of the type of data collected and analyzed for this study. Table 1 and 2 summarize the findings on students' reaction and the results of their self-efficacy measures. Table 3 and 4 display the results of our analysis of the focus group discussion and open ended questions in the reaction survey.

Students' feedback and self-efficacy

The analysis of students' feedback questionnaire (see Table 1) showed that they found ePlan^{TM} to be helpful (*M* = 4.17, *SD* = 1.03), informative (*M* = 4.17, *SD* = 1.27), and instructive in providing appropriate direction (*M* = 4.25, *SD* = 1.49) for the organizational analysis phase. They also indicated helpfulness of ePlan^{TM} to their analysis of the workforce environment, but indicated it was more informative than helpful (*M* = 3.67, *SD* = 1.23 for being helpful; *M* = 4.17, *SD* = 1.03 for being informative; *M* = 3.92, *SD* = 1.44 for being instructive). In formulating their approach to cause and gap analysis, ePlan^{TM} students also reported that it was helpful (*M* = 4.08, *SD* = 1.38), that it assisted students to prepare the issues to explore in this phase (*M* = 4.08, *SD* = 1.38), and that it guided them in how to analyze the performance gaps and their causes (*M* = 4.17, *SD* = 1.47). But when it came to ePlan^{TM}'s contribution to their identification and selection of solutions for performance problems, students did not rate it as highly as in other phases of the project (*M* = 3.67, *SD* = 1.72 for helpfulness; *M* = 3.5, *SD* = 1.68 for being informative; *M* = 3.75, *SD* = 1.66 for guidance the analysis).

Table 1. *Students' feedback on the usefulness of the EPSS for the phase of their team project*

PSA Phases	Feedback Statements	Percentage of Ratings						M	SD
		Not at all				Very much			
		1	2	3	4	5	6		
Organizational Analysis	Helpful in formulating the team approach	8.3	16.7	0.0	25.0	0.0	50.0	4.17	1.03
	Informative in preparing issues to explore	0.0	16.7	16.7	0.0	0.0	66.7	4.17	1.27
	Instructive in providing the appropriate direction	8.3	8.3	0.0	33.3	33.3	16.7	4.25	1.49
Work Environment Analysis	Helpful in formulating the team approach	8.3	8.3	16.7	41.7	25.0	0.0	3.67	1.23
	Informative in preparing issues to explore	8.3	16.7	0.0	25.0	50.0	0.0	4.17	1.03
	Guiding how to analyze the issues	8.3	8.3	16.7	25.0	33.3	8.3	3.92	1.44
Gap and Cause Analysis	Helpful in formulating the team approach	8.3	0.0	25.0	16.7	41.7	8.3	4.08	1.38
	Informative in preparing issues to explore	8.3	0.0	25.0	16.7	41.7	8.3	4.08	1.38
	Guiding how to analyze gaps & causes	8.3	0.0	25.0	16.7	33.3	16.7	4.17	1.47
Solution ID & Selection	Helpful in formulating the team approach	16.7	8.3	16.7	25.0	16.7	16.7	3.67	1.72
	Informative in preparing issues to explore	16.7	8.3	25.0	25.0	8.3	16.7	3.5	1.68
	Guiding how to analyze the issues	8.3	16.7	16.7	33.3	0.0	25.0	3.75	1.66

Students reported a significant improvement in their self-efficacy regarding the uses of the EPSS and conducting the phases of a PSA project (see Table 2). At the end of this experience, they found themselves very confident in using an EPSS in support of their future projects ($t[10] = 3.96$, $p = .003$). They also reported significant self confidence in conducting PSA projects successfully ($t[10] = 2.34$, $p = .04$), and performing an organizational/environmental analysis ($t[10] = 3.01$, $p = .01$). Students were also confident that they can successfully identify and select appropriate performance solutions ($t[10] = 2.56$, $p = .03$). The improvement in their self-efficacy reported for conducting cause and gap analysis was lower than the other phases and neither it was statistically significant

Table 2. *Difference in Reported Students' Self Efficacy before and After Using of EPSS*

Tasks	Pretest		Posttest		Difference	*t*	*df*	*p*
	Mean	SD	Mean	SD				
Use of EPSS to support PSA projects	63.64	24.61	90.91	7.0	27.27*	3.96	10	0.003*
Conduct PSA projects successfully	82.73	14.90	92.73	10.09	10.00*	2.34	10	0.041*
Conduct organizational / environmental analysis	72.73	22.84	94.55	8.2	21.82*	3.01	10	0.013*
Conduct gap & cause analysis	79.09	23.00	93.64	9.24	14.55	1.84	10	0.096
Select and recommend appropriate performance solutions	75.45	22.07	93.64	8.09	18.18*	2.56	10	0.029*

* $p < 0.05$

Students' discussion of the experience

Table 3 displays the statements inducted from students' comments made during the focus group discussion of their experience and in response to open-ended questions in the Feedback survey. The number of references made to each one of these issues is also shown in this table. Students most frequently (22 & 21) referred to the contribution of ePlan^{TM} in channeling their thoughts throughout the process and structuring their thoughts by providing them with a framework and a context. Further down in this ranking were the students' references to ePlan^{TM} helping the novices by clarifying the issues and providing a sense of reality for practicing PSA. Other positive references were made to the use of ePlan^{TM} for activities such as marketing PSA, providing opportunities for reflection and restatement of objectives, and facilitating transferring technical data to clients.

Table 3. *Categorization of students' comments and their frequency of references*

Use of ePlan^{TM} in PSA Course	**Freq. of References**
Channels your thoughts throughout the process	22
Structures your thoughts by providing a framework and context	21
Helps the novices by clarifying the PSA issues.	5
Provides a sense of realities of practicing PSA as profession	4
A bank of cumulative experiences on PSA/HPT.	2
A marketing tool for performance technology.	2
Provides instant access to tools & instruments	2
Complements class activities and content by providing an opportunity for their application.	2
Provides opportunities for reflection and restatement of objectives	2
Confirms your conceptualization of the process and analysis	2
Works like a KMS (Knowledge Management System)	1
Facilitates transfer of technical data to clients	1
Provides procedural guidelines for efforts on tasks throughout the phases of the project.	1
Facilitates team building in group projects	1

Students' negative comments were directed toward the design of ePlan^{TM} , its instructional design problems, and its interaction with the course layout. Table 4 displays the number of references students made to these issues. Out of a total of 87 negative references deducted from student comments, 16.1% stated that the software is not intuitive and 11.5% criticized its awkward navigation. Not being user friendly was listed on top in the instructional design problems category, including 16.1% of references. Redundant data entry requirements, lacking declarative and predictable sequence of events, and having too much text on the screen, were among other issues referred to more frequently in this category. For the course related problem category of comments, lack of sync and alignment with the course was the number one issue that students referred to (22.1%). They also reported that ePlan^{TM} is not appropriate for every PSA project (10.3%). A few references were also made (5.7%), on the inhibitive function of ePlan^{TM} that would not allow including other existing theoretical paradigms.

Table 4. *Categorization and frequencies of students' comments on the technical aspects of ePlan™*

Technical Difficulties with ePlan™	**Freq. and % of References**	
	Freq.	**%**
EPSS Design Problems:		
• Not intuitive interface	14	16.1
• Awkward Navigation	10	11.5
Instructional Design problems:		
• Not user-friendly Content/functions	14	16.1
• Redundant data entry requirements	7	8.0
• Presentation of PSA phases lacked declarative and predictable sequencing of events	3	3.4
• Too much text	3	3.4
• Unclear instructions on how to modify or what to fill in	2	2.3
Course Related Problems:		
• Lack of sync and alignment with the course schedule	20	22.1
• Not relevant or appropriate for every PSA project.	9	10.3
• Inhibitive to thinking out of the box to include other existing theoretical paradigms	5	5.7
Total number of References	**87**	**100%**

Discussion and Conclusions

The PSA course that provided the context for this experience was designed to accommodate for the current shift in instructional design programs toward the performance technology. Students in this course are interested in learning how to conduct performance analysis as a complementing skill to their instructional design competencies. Using a performance support system is what performance analysts may recommend to assist clients for improving performance. The use of the EPSS provided the performance consultants themselves with a tool to enhance their own performance in a course designed as a situated learning process.

Conducting PSA is a multifaceted task that requires a significant amount of concentration and planning. The analysts usually experience several issues at one time and perform multiple tasks simultaneously. In their contacts with the performance workforce for the purpose of observation and analysis, they have to make note of the causes, gaps, and solutions, even though they may have planned to consider those issues at a different visit. This nonlinear activity requires a large amount of forethought in planning an approach, exploring the right issues, reflection on implications of their processes, and taking an appropriate direction. According to the data collected at different points of this study, students found their use of EPSS to be supporting these very issues. In different phases of using ePlan™, students found it to be helpful to formulate their team approach, informative in preparing their exploratory issues, and instructive in guiding them through the process. The overall result of this experience was that ePlan™ assisted students with their self regulation required for their success in accomplishing the course objectives.

This PSA course happens to be one of the major courses of the Instructional Systems program at FSU. The course's practical approach to teaching PSA and its apprenticeship requirements are designed to prepare students for the real world of HPT and consulting. Because of these reasons, students may need additional confidence builder mechanisms to complement their practical experiences. The use of the EPSS proved to be one of those tools that provided the boost of confidence for students. At the end of their application of this EPSS, students indicated a significant improvement in their self-efficacy in using an EPSS for their future projects. They also reported an enhanced self-efficacy in conducting a PSA project on their own including its various phases. Gaining this type of rightful self image and practical experiences will contribute to students' development of competencies that assist them in working and succeeding in the real environment.

These findings were supported by students' comments in their discussion of the process. Based on their indications, the use of EPSS provided a positive learning experience, provided them with the right framework for their activities, and assisted with keeping them on track while performing a multifaceted task. Obviously, the use of an EPSS that becomes so important in performing a task must be instructionally sound. According to students' comments on the features and functions of ePlan™, one can summarize that an EPSS should truly support the performance, rather than hinder the effort of the user. Issues such as the intuitiveness of the interface, smooth navigation, user friendliness, and non-redundant data entry requirements should not inhibit the user if the purpose of the EPSS is to enhance users' performance or aid their learning.

In summary, the use of the EPSS in this study proved to be a positive experience for the users. They used it successfully to perform the tasks they were assigned. Part of this was because this particular EPSS was specifically designed for the type of use that was the subject of the course and for users like these students. In other words, this EPSS was tailor-maid for this professional population and therefore students found it rewarding to use because it was beneficial for practically managing their projects. The authors think this is an important issue to note for future developers of EPSSs. We also recommend that there should be an accurate alignment between the features of the software and the context in which is presented. In our case the course activities might have been better aligned with the phase of ePlan™ for its better success.

Nevertheless, we viewed the students' reports of using the EPSS as an enhancement to students' understanding of the HPT process by causing them to process and apply the knowledge and skills they acquired in other course activities. Finally, it seemed to enhance their self-efficacy by aiding them to monitor, record, reflect upon, and evaluate their own performance in the completion of students' class projects. We believe the addition of ePlan™ to the PSA course's activities can, with refinements to the EPSS and adjustments to the design of the course, can be a valuable asset for students' development of their knowledge and skills, and can support their self-regulated learning of performance analysis.

References

Ackoff, R. L. (1999). *Ackoff's best*. New York: John Wiley & Sons Inc.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248-287.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586-598).

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.

Dean, P. J. (1999). Designing better organizations with human performance technology and organization development. In H. D. Stolovitch and E. J. Keeps (Eds.), *Handbook of human performance technology*, 2nd ed. (pp. 321-334). San Francisco: Jossey-Bass Pfeiffer.

Gagné, R. M. (1985). What should a performance improvements professional know and do? *Performance and Instruction, 24*(7), 6-7.

Gagné, R. M., Briggs, L. J., & Wager, W. W. (1992). *Principles of instructional design*, 4th ed. Orlando, FL: Harcourt Brace Jovanovich.

Gagné, R. M., & Medsker, K. L. (1996). *The conditions of learning: Training applications*. Fort Worth, TX: Harcourt Brace College Publishers.

Gharajedaghi, J. (1999). *Systems thinking: Managing chaos and complexity*. Boston: Butterworth-Heinemann.

Gilbert, T. F. (1996). *Human competence: Engineering worthy performance*. Amherst, MA, and Washington, DC: Human Resources Development Press and International Society for Performance Improvement.

Hale, J. (1998). The performance consultant's field book, San Francisco, CA: Jossey-Bass

Harless, J. H. (1975). *An ounce of analysis is worth a pound of objectives*. Newman, GA: Harless

Performance Guild.
ISPI (n.d.). International Society for Performance Improvement. Retrieved from http://www.ispi.org on September 28, 2003.
Mager, R. F., & Pipe, P. (1997). *Analyzing performance problems*. Atlanta, GA: Center for Effective Performance Inc.
Robinson, D. G., & Robinson, J. C. (1995). *Performance consulting: Moving beyond training*. San Francisco, CA: Berrett-Koehler Publishers, Inc.
Rosenberg, M. J. (1990). Performance technology: Working the system. *Training*, 27(2), 42-48.
Rossett, A. (1999). Analysis of human performance technology. In H. D. Stolovitch and E. J. Keeps (Eds.), *Handbook of human performance technology*, 2nd ed. (pp. 139-162). San Francisco: Jossey-Bass Pfeiffer.
Rummler, G. A. (1999). Transforming organization through human performance technology. In H. D. Stolovitch and E. J. Keeps (Eds.), *Handbook of human performance technology,* 2nd ed. (pp. 47-66). San Francisco: Jossey-Bass Pfeiffer.
Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical skills*. Englewood Cliffs, NJ: Educational Technology Publications.
Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: from teaching to self-reflective practice* (pp. 1-19). New York: Guilford Press.

The Past, Present, and Future of Research in Distance Education: Results of a Content Analysis

Marcy P. Driscoll
Youngmin Lee
Florida State University

Abstract

The purpose of the study was to identify research trends in distance education through a content analysis conducted of 383 research articles published in four prominent distance education journals. Both topic and research method were coded and analyzed. An analysis of research topics indicated that approximately 102 (27%) were classified as design topics, 33 (9%) as development topics, 45 (11%) as management topics, 47 (12%) as evaluation topics, 38 (10%) as institution and operation topics, and 118 (31%) as theory and research topics. Results of coding research method revealed that 78 (20%) out of the 383 published in four journals focused on the theoretical inquiry, 47 (12%) on the experimental research, 138 (36%) on the case study, 23 (6%) on the evaluation research, 26 (7%) on the developmental research, and48 (13%) on the survey research. 23 (6%) were implemented with combined-inquiries. Result was discussed in terms of implications for future research in the field.

Introduction

The overall purpose of this study was to investigate research trends and issues in the field of distance education during the last six years (1997-2002). Many researchers know that a broad understanding of research in terms of topic and issues investigated previously can be very useful to advance the field (Saba, 2000; Serry, 1996). However, previous reviews do not seem to show a broad picture of research trends and issues in distance education.

In addition, the "state-of-the-art" coding scheme taken as the indicator for measuring and organizing the types of research topics and methods has not been quite exquisite. It is not so easily understood how it was established and how the topics were classified into the coding categories (Bauer, 2000; Krippendorff, 1980). Also, it is not sufficient to suppose which topics are more concerned and discussed (Szabo and Rourke, 2002). Therefore, new study is needed to exam the trends and issues in the field of distance education.

The main research questions consisted of: (1) What general research topics have been studied in distance education research; (2) What specific topics have been discussed in distance education research; and (3) Which methods have been applied and are prevalent in distance education research. Once these research topics, methods, and influential studies are identified, they can be used to review current research trends and explore new research directions.

Research design

A content analysis was conducted in order to analyze the trends and issues of distance education research. This method was chosen because it makes replicable and valid inferences from data to their context as well as describes trends in content (Krippendorff, 1980). The data sources of this study were all articles published from 1997 to 2002 in four journals. Exceptions are editorial, commentary, book review, and dialogue. The total number of selected articles was 383.

The four journals in the field of distance education are that were chosen for analysis as follow: (1) American Journal of Distance Education (AJDE); (2) Journal of Distance Education (JDE); (3) Distance Education (DE); and (4) Journal of Open and Distance Learning (JODL: Open Learning). These journals were selected based on their recognition as prominent journals among researchers and also having been previously chosen as data sources in earlier studies.

A new categorization system for encompassing distance education was developed to classify the articles published in terms of their topics. It consists of institutional/operational, design, development, management, evaluation, and theory/research topics.

- Design-related topics: need assessment, course scheduling, course design, instructional strategy development, course material design, and visual design.
- Development-related topics: course support system and material development, web-based learning management system building, online tools development, and online testing system development.
- Management-related topics: learning resource management, troubleshooting, attrition rate, faculty and staff support, learner support, and technical support.
- Evaluation-related topics: program quality control, assessment of learning outcomes, benefits and

cost analysis, Return on Investment (ROI), evaluation of supporting system, and peer feedback.

- Institutional and operational-related topics: administration, academic affairs, accreditation, certification, policy, payment, and budgeting.
- Theory and research-related topics: distance education theory building, review of literature, introduction to new research method, culture and gender issues, learning style, history of distance education and copyright law.

Three prominent keywords were also selected to identify detailed topics frequently adopted for the studies. Additionally, a new research method classification system has been developed. It includes theoretical inquiry, experimental research, case study, evaluation research, developmental research, survey research, and combined inquiries.

- Theoretical Inquiry: Theoretical review of literature and conceptual study for proposing new ideas in distance education.
- Experimental research: A study examining the effect of independent variable on a dependent variable.
- Case study: A study aimed at investigating a single individual, group, program, or organization qualitatively.
- Evaluation research: A study aimed at determining the impact of a project, program, model or software.
- Developmental research: A study aimed at designing, developing, and evaluating an existing or newly developed model, process, product, or technique.
- Survey research: A study addressing the distribution and return of responses in non-experimental situation.
- Combination of inquiries: A study synthesizing two or more research methods.

Experimental research was further classified to identify the types of statistical methods to be conducted. A coding scheme for the classification stemmed from the open-ended and emergent design approach, which means that the coding scheme were not specified in advance to the classification but rather constructed during actual data collection (Patton, 2002).

Results of the study

An analysis of research topics indicated that approximately 102 (27%) were classified as design topics, 33 (9%) as development topics, 45 (11%) as management topics, 47 (12%) as evaluation topics, 38 (10%) as institution and operation topics, and 118 (31%) as theory and research topics. As seen in Table 1, the theory and research topics, such as theory and model building, review of literature, and culture issues, were most likely chosen and researched. Subsequently, design and development research topics were used to report a course design, instructional strategy development, course material design, course support system, online material development, and online testing system development.

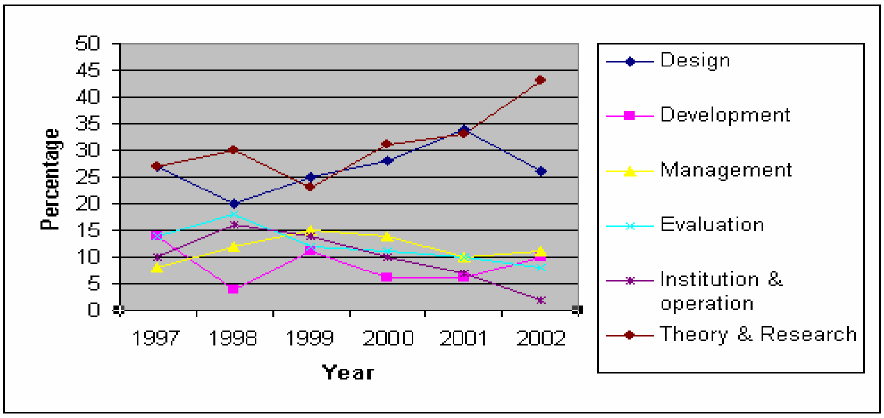
This result corresponds to that of other studies. The study of Anglin and Morrison (2000) showed that over 70% of the articles were classified as primary research and conceptual and theoretical studies. The study of Berge and Mrozowski (2001) presents that over 75% of the articles were associated with descriptive research topics. The study of Koble and Bunke (1997) also classified a large number of studies as theory and student characteristics issues.

Table 1. *Distance education research topics*

Type / Year	1997	1998	1999	2000	2001	2002	Total
Design	19	14	16	18	21	14	102 (27%)
Development	10	3	7	4	4	5	33 (9%)
Management	6	8	10	9	6	6	45 (11%)
Evaluation	10	12	8	7	6	4	47 (12%)
Institution & Operation	7	11	9	6	4	1	38 (10%)
Theory & Research	19	21	15	20	20	23	118 (31%)
Total	**71**	**69**	**65**	**64**	**61**	**53**	**383 (100%)**

*In the Paranthesis is the percentage of the research topic in that year

Figure 1. Percentage of research topic by publication year



Three prominent keywords were selected from each article by the year. The statistics show that 213 keywords in 1997, 207 keywords in 1998, 195 keywords in 1999, 192 keywords in 2000, 183 keywords in 2001, and 159 keywords in 2002 were identified and classified by the year. As seen in Table 2, specific topics include (1) an interaction, program evaluation, collaboration, and video conferencing in 1997; (2) a quality of the program, learners' perception, effectiveness of a program, and faculty support in 1998; (3) learners' perception, video conferencing, learners' participation, and learners' attrition in 1999; (4) A collaboration, learners' perception, video conferencing, effectiveness of a program, and learner's achievement in 2000; (5) A cross-cultural issues, faculty support, video conferencing, collaboration, and online learning barriers in 2001; and (6) Problem based learning (PBL), interaction, learners' attitude, learners' perception, flexible learning, learners' satisfaction toward a program, and online tutor in 2002.

Table 2. *Distance education research specific topics by the keyword*

	1997	**1998**	**1999**	**2000**	**2001**	**2002**
1	Interaction (12)	Program quality (8)	Learners' perception (5)	Collaboration (10)	Cross-cultural issues (11)	PBL* (8)
2	Program evaluation (8)	Learners' Perception (7)	Video conferencing (4)	Learners' Perception (8)	Faculty support (5)	Interaction (7)
3	Collaboration (8)	Program's effectiveness (5)	Learners' Participation (4)	Video conferencing (5)	Video conferencing (4)	Learners' attitude (6)
4	Video conferencing (7)	Faculty support (5)	Learners' Attrition (4)	Program's effectiveness (4)	Collaboration (4)	Learners' perception (4)
5	Learning outcome (6)	Interaction (4)	Collaboration (3)	Learners' achievement (4)	Barrier to online learning (3)	Flexible learning (4)
6	Online learning model (5)	Instructor's Leadership (4)	Program evaluation (3)	Need analysis (4)	Assessment of outcomes (3)	Learner's satisfaction (4)
7	Program quality (4)	Theory development (4)	Metacognition (3)	Program evaluation (4)	Interaction (3)	Tutor (4)
8	Faculty support (4)	Institution issues (4)	Learners' achievement (3)	Faculty support (3)	Learners' performance (3)	Program evaluation (3)
9	Learners' Perception (3)	Leaning strategy (3)	Faculty support (3)	SDL* (3)	Program's quality (3)	Faculty support (3)
10	Learner's Satisfaction (3)	Assessment of outcomes(3)	Tutor (3)	Cost issues (3)	Learners' persistence (3)	Scaffolding (3)

*In the paranthesis is the frequency of the keyword in that year
*SDL: Self-Directed Learning PBL: Problem Based Learning

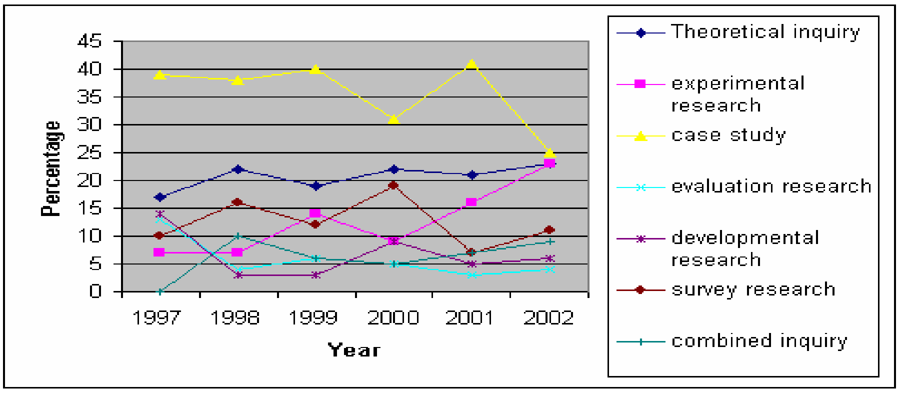
The table indicates that an interaction, perception, collaboration, conferencing, evaluation, and support show an increasing number of attention throughout the years. These keywords were included, and closely related, with the theory, research topics, and the design topics in Table 1. The reason why the cross-cultural issue ranked first in 2001 is that the Distance Education journal 22(1) was a special edition for the cross culture issues in distance education. In most cases, however, cultural issues in distance education have not been prioritized by the researchers in the field.

Results of coding research method revealed that 78 (20%) out of the 383 published in four journals focused on the theoretical inquiry, 47 (12%) on the experimental research, 138 (36%) on the case study, 23 (6%) on the evaluation research, 26 (7%) on the developmental research, and48 (13%) on the survey research. 23 (6%) were implemented with combined-inquiries. As indicated in Table 3, the case study was the most frequently used method in distance education. This research method was followed by theoretical inquiry and survey research. These studies showed that one-third of these were classified as case studies. The purpose of these groups of studies and research was to focus on investigating single individual, group, program, or organizations qualitatively. These studies have led researchers to examining a particular group and program by using the review of the literature, interviews, field observation, and focus group meeting.

Table 3. *Distance education research method*

Type / Year	1997	1998	1999	2000	2001	2002	Total
Theoretical Inquiry	12	15	12	14	13	12	78 (20%)
Experimental research	5	5	9	6	10	12	47 (12%)
Case study	28	26	26	20	25	13	138 (36%)
Evaluation research	9	3	4	3	2	2	23 (6%)
Developmental research	10	2	2	6	3	3	26 (7%)
Survey research	7	11	8	12	4	6	48 (13%)
Combined inquiries	0	7	4	3	4	5	23 (6%)
Total	**71**	**69**	**65**	**64**	**61**	**53**	**383 (100%)**

Figure 2. Percentage of research method by publication year



These studies indicate that case studies were used on one-third of them. The purpose of these many studies was to focus on investigating single individual, group, program, or organization qualitatively. These studies lead the researchers to examining a particular group and program by using the review of the literature, interviews, field observation, and focus group meeting. It is followed by theoretical inquiry, survey research, and experimental research. Although the results of Berge and Mrozowski (2001) reported two-thirds of their studies as descriptive statistics followed by case studies, the results of them are not far different than the result of this study. If theoretical inquiry, evaluation research, development research, and survey were to be incorporated into descriptive study, a good portion of these categories became pertinent to the study of Berge and Mrozowski.

Experimental studies were further specified to report the scope of statistical methods that researchers in the field have applied over the years. However, the result of this statistical method has a limitation before generalizing it because the number of experimental studies is relatively small. The result conveyed that the number of studies using statistical methods is increasing through the years. The researchers in the field frequently have used ANOVA (Analysis of varinace), regression, chi-squre, and factor analysis as the main methods. In particular, factor analysis has been applied to the studies almost every year. This method was used to indentify psychological elements such as persistence, acceptance, perception and their relationship in distance education.

Table 4. *Frequency of experimental researches by statistical method*

Type / Year	1997	1998	1999	2000	2001	2002	Total
T-test	1	1	2				4 (9%)
ANOVA, ANCOVA	1	1	2		1	3	8 (17%)
Regression	2		1		3	2	8 (17%)
Chi-square	1	1		2	3	1	8 (17%)
Correlation			1	1	1	2	5 (11%)
MANOVA, MANCOVA						1	1 (2%)
Factor analysis		1	1	2	2	2	8 (17%)
Cluster analysis			1				1 (2%)
Path analysis		1	1				2 (4%)
Meta analysis				1		1	2 (4%)
Total	**5**	**5**	**9**	**6**	**10**	**12**	**47 (100%)**

Discussion and future remarks

The current results of the study convey certain implications on which future research in the field of distance education can expand. First, developmental research approach (Richey & Nelson, 1996) has been rarely found in distance education journals (Driscoll & Dick, 1999). Almost 30% of articles published have been classified into the case studies to investigate a single individual, group, program, or organization qualitatively. However, many studies primarily focus on describing the context of a project, process, and final output, rather than the description of unique theory supporting the purpose of the study. Thus, it is relatively difficult to understand what theory is implicated on the design and development of the project as well as on the revision of existing theory adopted and construction of new theory. It is one of many limitations in the case study that it may be useful only to the researchers and practitioners involving similar project (Anglin & Morrision, 2000).

Given the comparatively small number of theory-based studies in the field of distance education, why there have not been more theory-based studies remains unanswered. Part of the problem may be a lack of theory-driven research methodology for distance education. Holmberg (cited from Simonson, Smalidino, Albright, & Zvacek, 2003) described that:

"Distance education has been characterized by a ***trial-and-error approach*** with little consideration being given to a theoretical basis for decision making…The theoretical underpinning of distance education are fragile. Most efforts in this field have been practical or mechanical and have concentrated on the logistics of the enterprise"

In addition, the lack of developmental researches may also be due to which most practitioners have preferences solely towards the result of the study. That is, researchers may think that only the best practice as the result of a project will reduce the threats to a failure of implementation of new distance learning program as well as encourage the practitioners to achieve similar results of the original program. In so doing, they have been more interested in reporting the technical processes, practical strategies, and statistical results. There is no doubt that the user-centered research is useful and effective to the practitioner's works. However, at the end, the researchers as well as practitioners may fail to take a notice of all theories comprehensively in results which trial-and-error symptom continues repeatedly.

References

Anglin, G. J. & Morrison, G. R. (2000). An analysis of distance education research: Implications for the instructional technologist. *Quarterly Review of Distance Education, 1*,189-194.

Bauer, M. (2000). Classical content analysis: A review. In M. Bauer & G. Gaskell (Eds.), *Qualitative Researching with Text, Image and Sound* (pp. 131-151). Thousand Oaks, CA: Sage.

Berge, Z. L., & Mrozoski, S. (2001). Review of research in distance education, 1990-1999. *American Journal of Distance Education, 15*(3), 1-15.

Driscoll, M., & Dick, W. (1999). New research paradigms in instructional technology: An inquiry. *Educational Technology Research and Development, 47*(2), 7-18.

Koble, M., & Bunker, E. (1997). Trends in research and practice: An examination of The American Journal of Distance Education. *American Journal of Distance Education, 11*(2), 19-38.

Krippendorff, K. (1980). *Content analysis: an introduction to its methodology*. Beverly hills, London: Sage publication.

Patton, M. Q. (2002). *Qualitative Evaluation and Research Methods*. (3rd ed.). Newbury Park, CA: Sage.

Richey, R.T. & Nelson, W.A. (1996). Development research. In D. Jonassen (Ed)., *Handbook on*

educational communications and technology (pp.1230-1232). London: MacMillan.
Rourke, L., & Szabo, M. (2002). A content analysis of the Journal of Distance Education 1986-2001. *Journal of Distance Education, 17*(2), 63-74.
Saba, F. (2000). Research in distance education: A status report. *International Review of Research in Distance Education, 1*, 16-23.
Sherry, L. (1995). Issues in distance learning. *International Journal of Educational Telecommunications,*
1(4), 337-365.
Simonson, M., Smaldino, S., Albright, M., & Zvacek, S. (2003). *Teaching and learning at a distance: foundations of distance education* (2nd ed.). Upper saddle river, NJ: Prentice Hall.

Expertise in Complex Problem Solving

Deniz Eseryel
Syracuse University

Introduction

Researchers in the field of problem solving have long been struggled by the absence of universally agreed-upon definitions of many of the basic terms used (e.g., Smith, 1991). Among these basic terms are expert, novice, problem, and even problem solving. Consider, for instance, the following most-commonly cited definitions of a problem:

- A problem exists when the goal that is sought is not directly attainable by the performance of a simple act available in the animal's repertory; the solution calls for either a novel action or a new integration of available actions. (Thorndike, 1898, cited by Sheerer, 1963, p.18)
- A problem occurs…if an organism follows a goal and does not know how to reach it. Whenever a given state cannot be converted into the desired state directly through the activation of obvious operations, thinking is needed to construct mediating actions. (Duncker, 1935, translated and cited by Frensch & Funke, 1995, p.6)
- A question for which there is at the moment no answer is a problem. (Skinner, 1966, p.225)
- A problem is a stimulus situation for which an organism does not have a ready response. (Davis, 1973, p.12)
- A problem is a "stimulus situation for which an organism does not have a response,"…a problem arises "when the individual cannot immediately and effectively respond to the situation." (Woods, Crow, Hoffman, & Wright, 1985, p.1)
- A person is confronted with a problem when he wants something and does not know immediately what series of actions he can perform to get it. (Newell & Simon, 1972, p.72)
- Whenever there is a gap between where you are now and where you want to be, and you do not know how to find a way to cross that gap, you have a problem. (Hayes, 1980, p.i)
- A problem is an unknown. That is, if we have a goal and do not know how to reach that goal, there is an unknown, so we have a problem. …finding the unknown must have some social, cultural, or intellectual value to someone. If no one believes that it is worth finding the unknown, there is no perceived problem. (Jonassen, 1997, p.109)

The definitions of a problem differ primarily on three dimensions. First, they differ in terms of their semantic content. That is, the focus on either the absence of a task relevant *response* or a task-relevant *thought*. Second, the definitions differ in how fuzzy their boundaries are. And finally, the definitions differ in terms of their category size, that is, how many tasks can be classified as problems. Nevertheless, most definitions do appear to share an important attribute, namely, the focus on the *gap* between the capabilities of the problem-solver and the requirements of the task at hand. In other words, a problem is said to exist only if there is a barrier between the state given in the current situation and the goal state in the head of the problem solver. So, problems are not task-specific. A problem is not defined by the features of the task. Rather, a problem exists or do not exist depending on the interaction between task requirements and the problem solver. A certain task may constitute a problem for one solver, but may not for another. On the contrary, task-focused definitions lead that a given task either constitutes or does not constitute a problem for all solvers.

What is, then, problem solving? Again, there is no agreed-upon definition for problem solving. Consider the following most-commonly cited definitions:

- Problem solving is defined as any goal-directed sequence of cognitive operations. (Anderson, 1980, p.257)
- …problem solving is defined here as a goal-directed sequence of cognitive and affective operations as well as behavioral responses for the purpose of adapting to internal or external demands or challenges. (Heppner & Krauskopf, 1987, p.375)
- What you do, when you don't know what to do. (Wheatley, 1984, p.1)
- A problem is an unknown. That is, if we have a goal and do not know how to reach that goal, there is an unknown, so we have a problem. …finding the unknown is the process of problem solving. (Jonassen, 1997, p.109)

As in the case of problem solving, the existing definitions differ primarily on (i) semantic content; (ii) category size, and (iii) how fuzzy their boundaries are. Information processing view of problem solving suggests that problem solving requires the mental representation of the situation, known as the problem space (Newell &

Simon, 1972). The process of solving a problem, then, requires some active manipulation of this problem space. The success of the problem solver is highly affected by how well the mental representation of the situation reflects the *actual* situation. In other words, how well the constructs and the relationships between the constructs of the problem space resemble or fail to resemble to those of the actual situation.

One of the difficulties associated with problem solving is that not all problems are created equal. There are problems which can be solved with a few mental steps and there are problems that require extensive thinking. There are problems that we have never encountered before and there are problems that we are familiar with. There are problems which have clear goals and there are problems which the goal is not that clear to us. Hence, problems can be distinguished on any number of different dimensions. Imagine your typical 5-year old sliding all over the room trying to build the world's largest building using her Lego building blocks. And imagine the researcher trying to find a cure to AIDS epidemic. Both the 5-year-old and the seasoned researcher engage in complex problem solving (henceforth CPS).

Humans have difficulty in understanding and in making robust policies concerning complex domains (Dörner, 1996). There are a number of reasons for this difficulty (Dörner & Wearing, 1995; Funke, 1991). First of all, any given complex problem solving situation may involve multiple goals and it is very difficult to define goals *operationally*. Often, this requires decomposing the 'global' goal into many subgoals but this leads to another difficulty: As time is always limited it is necessary, not only for one action to serve more than one goal, but also to order the priority of these goals. However, as the most important and urgent goal is being addressed the variables in the system may interact in such a way that lead to the requirement of reconsidering the overall system goal (e.g., MacKinnon & Wearing, 1980, 1983). Nevertheless, in some cases, it may not be necessary to act at all to reach ones goals as the system's development produce the goal state independently. If, however, the system does not move autonomously in the desired direction, it is necessary to act, taking into account the autonomous developmental tendencies of the system (Frensch & Funke, 1995). In any event it is necessary but challenging to predict what will happen to the system, as some of the goals may be contradictory which require reasonable trade-offs.

Another series of difficulties arises due to the challenges associated with constructing mental representation of the system to be managed (e.g., Dodds, Lawrence, & Wearing, 1991; Markman, 1998; Putz-Osterloh, & Lemme, 1987; Sonenberg, Lawrence, & Zelcer, 1992). A mental representation is necessary in order to predict the direct and immediate as well as the side and long-term effects of the decisions. It is however challenging due to the intransparency of the complex domains. Complex domains typically involve large number of variables with high-degree of connectivity. Changes in one variable may affect a number of other variables, making it very difficult to anticipate all possible consequences of a given situation. Therefore, it is not possible to directly observe all of the variables involved or the relationships between them. Furthermore, not every action shows immediate consequences. Some of the effects may occur with time delay. All of these factors make it very challenging for building the mental model of the complex problem domain necessary to solve complex problems and decision-making.

Research on Expertise in Complex Problem Solving

It is possible to trace the studies on expertise back to the beginning of nineteenth century, although under a variety of different labels (e.g., Binet and Simon, 1905, cited in Sternberg, 1995). However, the modern study of expertise can be said to start with DeGroot (1965). Surprisingly, only a very few studies on expertise complex problem solving seriously address the question of just what exactly an expert is. For the most part, investigators operationally define expertise in terms of outstanding performance on the selected task. Sternberg (1995, p.297) identifies ten different views of expertise. These include:

- *General-Process View*: Experts are those who solve problems by different processes from those used by nonexperts or who utilize the same processes more efficiently than do nonexperts;
- *Quantity-of-Knowledge View*: Experts simply know more than nonexperts;
- *Knowledge-Organization View*: Experts organize their knowledge more effectively than do nonexperts;
- *Cognitive-Complexity View*: Experts are cognitively more complex in the information they can process, perhaps due to the superior working memory capacity;
- *Superior Analytical Ability*: Experts can use the knowledge they have more effectively than nonexperts;
- *Superior Creative Ability*: Experts can create new knowledge based on the existing knowledge they have more effectively than nonexperts;
- *Superior Automatization Ability*: Experts do things more adeptly and automatically than do nonexperts;
- *Superior Practical Ability*: Experts know the ropes much better than nonexperts, in other words they know how to get ahead in their field;
- *Expertise by Virtue*: An expert is an expert by virtue of being labeled as such;
- *Synthetic View*: Expertise is a prototype, and is rarely reached in its purely prototypical form. Rather, people have aspects of expertise, namely, the other nine views mentioned above.

American & European Views of Expertise in Complex Problem Solving

There are significant differences in how researchers in United States and Europe are going about the research on expertise in complex problem solving (Table 1).

In United States, the research on expertise in complex problem solving can be traced back to the pioneering work of Newell, Shaw, and Simon (1958) and Miller, Galanter, and Pribram (1960), which shifted much of the emphasis in American experimental psychology from stimulus-response theories of behavior to cognitively based theories. However, the dominant contemporary American approach to complex problem solving originated in Europe by DeGroot (1965) on expert chess players, which was later extended by Chase and Simon (1973). The main impact of this work was to shift thinking about expertise from a point of view emphasizing general information processing to one emphasizing the preeminence of knowledge (Chase & Simon, 1973). Typical American studies have compared experts versus novices in the performance of tasks requiring high levels of professional expertise in technical domains. The view of expertise that has emerged from this line of research is "one of a problem solver who has memorized a huge number of tricks of the trade, rather than one who has developed a single powerful strategy for reasoning" (Hunt, 1991, p.394). In sum, "Americans have primarily studied expertise in knowledge-rich domains, and have attributed expertise to greater knowledge about these domains" (Sternberg, 1995, p.298). The American approach has a number of strengths. One is a tradition of careful experimental study. Another is the close observation of true experts at work. However, there is something nonprofound about showing that the difference between experts and novices is that experts know more than novices. Furthermore, American approach emphasized the intensive study of extremely specialized pursuits and it is hard to make the case that it generalizes to other pursuits. Also, it is possible that, in the American approach, the experts might be studied in do not necessarily render them experts. Finally, there has been a great attention focused on group differences between experts and novices in American approach, but relatively little attention on individual differences. Without considering individual differences, it is difficult to understand expertise.

Table 1. *Comparison of predominant American and European studies of expertise in complex problem solving*

	American Research	**European Research**
The View of Expertise	- Quantity-of-Knowledge - Knowledge-Organization	- Cognitive Complexity
Task Selection	- Problems Requiring High Levels of Professional Expertise in Knowledge-Rich Domains	- Everyday Problems
Approach	- Expert-Novice Differences	Task-Centered

Two principle approaches to studying expertise in complex problem solving are observed in European tradition. One approach was pioneered by Broadbent (1977) in England; the other was initiated by Dörner (1987) in Germany. The key difference between the two traditions is that in the former, one is able to specify a precise rule via a mathematical formula that would optimize problem solving, whereas in the latter, the problems are so complex that it is questionable whether we could ever devise any mathematical or even computer simulation that would clearly optimize performance. However, in both approaches the problems can be given to anyone of at least average intelligence and get them solve the problems. There are several advantageous features of the European approach: (i) Almost anyone can be a subject; (ii) problems are about life in general; (iii) the combination of British and German approaches look at problems that are so hopelessly complex and at problems whose optimal solution can be characterized by a mathematical function so that it is possible to compare people's performance to the optimal; (iv) the studies paid more attention to the individual differences. Despite the attractive general features, it is more difficult to specify any large set of principles of problem solving following European approach. Also, especially in German approach, it is very difficult to find a reliable and valid way of scoring performance in such complicated problems. Furthermore, the European approach is too task-centered to yield more general theories about problem solving.

In sum, American and European approaches are complimentary rather than contradictory. Both approaches enhance our understanding of complex problem solving.

Expertise in Instructional Design Problem Solving

In a recent study exploring the nature of expertise in complex problem solving, the author investigated the expert responses to complex instructional design problems. Instructional design is chosen as the test-bed for studying expertise in complex problem solving due to its compelling attributes as a dynamic complex domain.

The domain of instructional design is complex due to the dynamic interactions between the various dimensions of learners, learning situations, learning goals, and learning technologies involved (Spector & Anderson, 2000). Furthermore, instructional designers usually have to operate within many project constraints, which adds to the level of complexity. Most instructional situations call for ill-structured, poorly defined problem solving because they typically have vaguely defined goals, unstated constraints, multiple solution paths, no consensual agreement on the appropriate solution, and multiple criteria for evaluating solutions (Ertmer & Quinn, 1998; Goel & Pirolli, 1989; Jonassen, 1997; Perez, Johnson, & Emery, 1995; Spector & Anderson, 2000).

The goal of this research was to find out whether expert instructional designers exhibit recognizably similar patterns in their responses to complex instructional design scenarios. Operational definition for 'expert' included those who have at least a master's degree in the field and a minimum of ten years of hands-on experience (Chase and Simon, 1973; Chi et al., 1988; Goel & Pirolli, 1989; Perez et al., 1995; Simon, 1981). Additional criteria for selection of experts included recognition by peers and a graduate degree in a related field (Perez & Neiderman, 1992).

Participants

Six expert instructional designers participated in this study. All of them held graduate degrees (5 with Ph.D. and 1 with master's) in instructional design or a related field. They were employed in academia or in business. Mean age of the participants was 55.5 years. All of the participants had experiences in practicing instructional design and development; teaching ID and related courses; and conducting research in the field of instructional design and technology. The average for full-time experience in practicing instructional design and development was 19.83 years (min. 10 yrs; max. 34 yrs.); for teaching ID related courses, it was 5.33 years (min. 0 yr; max. 15 yrs); and for field research, it was 5.75 years (min. 1 yr; max. 18 yrs). Table 2 presents frequency counts of instructional design activities participants regularly performed in their current job.

Table 2. *Frequency distribution of instructional design activities participants regularly performed*

ID Activity	Frequency (N=6)	Percent
Conduct needs assessment	3	0.50
Determine solution alternatives and approaches	6	1.00
Propose solutions	4	0.67
Write learning objectives	4	0.67
Conduct task analysis	2	0.33
Identify types of learning outcomes	3	0.50
Assess learner's entry-level skills	3	0.50
Assess learner characteristics	5	0.83
Develop test items	3	0.50
Select instructional strategies	6	1.00
Select media formats	5	0.83
Conduct formative evaluations	4	0.67
Conduct summative evaluations	3	0.50
Manage instructional/training projects	4	0.67

Materials

Materials consisted of three parts: (i) A background survey; (ii) A worked out example including a sample problem scenario and a sample solution for that scenario; and (iii) Problem scenario to be worked out by participants.

The problem scenario was constructed in such a way that it resembled an authentic instructional design problem. It included design and development of instructional materials on the topic of integrating technology into middle school science teaching to be delivered via the Internet to rural teachers all across the USA within established project specifications.

Procedure

The study was conducted in two consecutive sessions. The first session was an informational session. Biographical information of participants was collected via background survey. Then, a brief description of the second session was presented.

During the second session, the participants were provided with an example explaining the task they were about to undertake. Then, they were given the problem scenario and were asked to: (i) take some time to individually reflect on the problem scenario and provide their assumptions and contextual remarks; (ii) create a representation of the given of the given problem scenario, identifying the key factors (such as entities, rates, or processes) and the relationships between the key factors; (iii) annotate each node (factor) and each link (relationship) in their representation by further elaborating on the details; (iv) provide recommendations for the solution of the given problem scenario based on their analysis as reflected in their representations.

Protocol analysis was used to analyze participants' annotated representations of given ID problem scenario, their assumptions and their recommendations for solution. Three main sources of information were used for the development of the coding scheme: (i) instructional design elements, which emerged from the initial qualitative analysis of participant data; (ii) established elements and activities in instructional design models (for a review of instructional design models, see Gustafson & Branch, 1997); and, (iii) general methods and guidelines suggested by Ericsson & Simon (1984) to develop a coding scheme for protocol analysis. Through frequency counts, the researcher sought for a recognizably similar pattern between the expert representations.

Results

The analysis of the protocols has consisted of two steps, namely: the examination of participant protocols in search of overall patterns and the coding of events related to those patterns.

The qualitative analysis was guided by the generation of initial research hypothesis to further elaborate on the similarities between the expert responses to the given instructional design problem. These questions evolved through the initial analysis of the protocols. The questions addressed how expert instructional designers approached complex design problem solving task and whether it was possible to graphically represent the common elements in their approach.

Table 3 presents the frequency distribution of resulting coding categories. From the ID process perspective, it was evident that expert instructional designers viewed the instructional design process in its entirety, including all of the design elements (i.e., instructional analysis, design, development, evaluation, and feedback) that are addressed in a generic ID model.

A number of common themes emerged from the expert protocols. These confirmed the findings of the previous studies (Goel & Pirolli, 1989; Perez et al., 1995; Rowland, 1992) on the characteristics of instructional design experts. Experts tried to get a holistic picture of the problem situation, viewed the problem in its entirety, and then they focused on one component of the problem, generated an initial design and evaluated it. Based on the results of evaluation, they accepted, rejected, or modified it. The expert designers then extended the particular solution to other areas of the "whole design" and evaluated it. From the responses of the participants, it was evident in this study that expert instructional designers viewed the instructional design process in its entirety, including all of the design elements (i.e., instructional analysis, design, development, evaluation, and feedback) that are addressed in a generic ID model.

Table 3. *Coding categories and their frequency counts*

ID Elements (Coding Categories)	Frequency (N=6)	Percent
1. Instructional analysis (includes determining goals and objective; task analysis; target group analysis; etc.)	2	0.33
2. Design (includes sequencing, determining modules, duration of instruction, construction of design templates, test items, etc.)	6	1.00
3. Develop module (includes instructional materials and media)	6	1.00
4. Pilot testing (includes formative evaluation of instructional materials developed, analysis of the pilot test results, and feeding the results of pilot test to improve design template)	5	1.00
5. Client (includes requirements, constraints, flexibility, budget, deadlines, satisfaction, etc.)	3	0.50
6. President (includes requirements, constraints, commitments, and flexibility)	3	0.50
7. Project budget (estimated budget vs. actual costs)	4	0.67
8. Personnel- requirements, expertise, availability (design & development team, SME, pilot test group, etc.)	5	0.83
9. Delivery schedule (estimated vs. actual)	5	0.83
10. Overall design & development time (estimated vs. actual)	5	0.83
11. Outsourcing	2	0.33
12. Team expertise	3	0.50
13. Quality of materials developed	3	0.50

Expert Causal Influence Diagram Guided by the research question, a second level of analysis aimed at constructing a causal influence diagram (CID) based on the similarities of the expert representations. This stage involved both the analysis of the nodes (factors) and the relationship between these nodes (links) across expert representations. Annotations of nodes and links were used to clarify the meaning attached to these elements by the experts.

Figure 1 presents the causal influence diagram developed based on the similar patterns between representations of expert instructional designers. In a way, this diagram represents the underlying model that expert instructional designers utilize while they are dealing with similar instructional design problems. According to this expert causal influence diagram (Figure 1), expert's mental model of the key factors and relationships of these key factors in the given problem case can be described as follows: Client's requirements, determines the project size, scope, budget, and timelines. Increased client requirements mean increased pressure on the personnel to meet client requirements. Initial project specifications state the number of personnel that can be hired for the project. As the size of the project increases, so does the required number of personnel. If more personnel are hired then it will be possible to more quickly design and develop the materials. Therefore, one of the recommendations from experts was to try to re-negotiate with the client & the president of the design company to hire additional staff or to outsource some of the design work.

More modules mean more pilot testing. Pilot testing insures the quality of the materials produced, but it also slows down the delivery of the materials. However, pilot testing also contributes to the team expertise. The more pilot testing is conducted, the more knowledge and experience they will have about what works, what doesn't work, and how to design the materials better and quicker. If team expertise increases, the need for more personnel decreases because existing personnel will work faster and better. More materials delivered means more client satisfaction. Another way to increase client satisfaction is to increase the quality of materials delivered. As client satisfaction increases there will be more orders for the instructional design company. In turn, pressure of the president on the personnel will decrease.

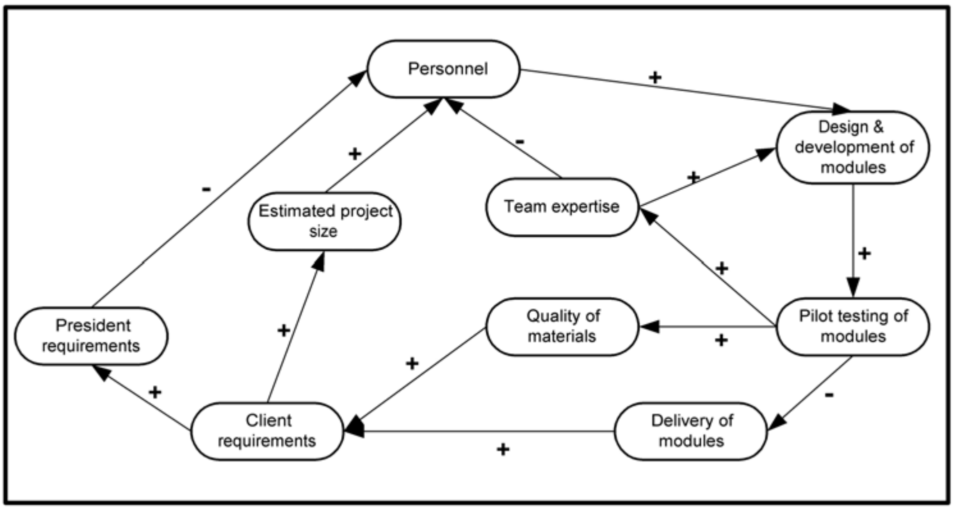


Figure 1. Casual influence diagram developed from the similar patterns in expert responses

Expert Recommendations for Solution As far as the recommendations for solution are concerned, a number of common themes emerged from the expert protocols. First, experts wanted to examine the whole production process, including resources (personnel, budget estimates, etc.), task division, and production schedule to determine causes of the problem. The second common expert reaction was to analyze the status of remaining resources and other production variables. Based on these two sources of input, experts typically focused on the main piece of the problem (according to their individual judgments), recommended a solution, and extended their solutions to include other pieces of the overall problem. Based on their earlier experiences, some experts provided a number of rule-of-thumbs (heuristics) that could be applied to the given problem scenario. After the early analysis of the problem, two main approaches were observed. Some of the experts wanted to report this to president and to the client for a possible renegotiation before revising the project management plan. Others did not consider this as an alternative. They preferred to revise the overall project management piece to ensure on time delivery. In order to push the production forward, a number of alternatives were proposed. The most common ones included construction of design templates based on the pilot testing data of earlier modules, reusability of previously developed materials, and outsourcing.

Team expertise also emerged as an important factor in the development process. Experts argued that after the development and pilot testing of the first four modules, the ID team would have the expertise to more quickly finish the remaining modules. The feedback from the pilot testing from the first four modules could be used to improve the design template. By this way, it would be possible to use this template to develop the rest of the modules. If time is of concern, experts suggested not to pilot test the remaining modules. They agreed on that it would be a compromise the quality of the materials, however, they argued that since previous materials were pilot tested and the design template was improved based on the results of the pilot testing it would not be a big compromise as far as the quality was concerned.

In this study, it was possible to identify a common expert model in response to the given ID problem scenario. Proposed dissertation study is aimed at validating these finding with larger number of experts and also comparing the expert model with those produced by novices. Also, the coding scheme developed from the pilot study will be useful for developing initial coding scheme for the proposed study. The pilot study also proved valuable in testing the instructional design task and the materials to be used for the proposed dissertation study.

Conclusion & Discussion

Although the study described in this paper focused on the similarities between the expert instructional designers, differences between expert representations were also observed in the data. Although, it is not possible to scientifically argue where these differences come from a hypothesis is called for. It is the hypothesis of this author that the different contexts in which the experts have to work could explain some of the differences between expert representations. Furthermore, the external representations of experts may vary to a greater extent in dependence on their epistemic beliefs concerning ID and ISD. For instance, consider the research group of there are some experts who would not apply a procedural ID-model such as the Dick-and-Carey-model but rather they work on the theory and research on mental models in order to create learning environments that improve the learners´ model building and revision. In a similar vein Mayer and others work insofar as the refer to the Dual Code theory of Paivio or Cognitive Load theory in order to construct or synthesize learning tasks. In these research camps, there is no involvement in the ID-models of the first and second generation. Certainly, these research camps are different with regard to the epistemic beliefs about learning and ID as a complex problem-solving task.

Therefore, this line of research could be furthered by distinguishing different types of experts in dependence on the concrete design problems they have to solve. The studies of Krems (1995) on cognitive flexibility in complex problem solving suggest that there are several "types of experts" in the field of ID such as (1) people who have to develop instructional programs, (2) scholars who teach to develop instructional programs, and (2) meta-theorists who develop conceptual models such as the Dick-and-Carey-model that doesn't fit the requirements of problem solving in the field of ID. It might be further helpful to classify the field of ID-models in accordance with three "generations": (a) the development of procedural and conceptual models of ID and ISD (i.e. the first generation), (b) attempts to automatize parts of the ID-process (the second generation: AID-models: Spector and others), and (c) theory and research-based approaches of ID such as Goal-based scenarios, Anchored Instruction of CTGV, and model-based learning and teaching (Gibbons, 2001; Milrad, Spector, & Davidsen, 2002; Seel, 2003).

References

Anderson, J. R. (1980). *Cognitive psychology and its implications*. New York: Freeman.

Broadbent, D. E. (1977). Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology*, 29, 181-201.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in dress. In W. G. Chase (Ed.), *Visual Information Processing* (pp. 215-218). New York: Academic Press.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Davis, R. B. (1985). Solving the "three switch" problem: A case study. *Journal of Mathematical Behavior*, 4, 281-291.

DeGroot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton.

Dodds, A. E., Lawrence, J. A., & Wearing, A. J. (1991). What makes nursing satisfying: A comparison of college students' and registered nurses' views. *Journal of Advanced Nursing*, 16, 741-753.

Dörner, D. (1987). On the difficulties people have in dealing with complexity. In J. Rasmussen, K. Duncan, & Leplat (Eds.), *New Technology and Human Error* (pp.97-109). New York: Wiley.

Dörner, D. (1996) (Translated by Rita and Robert Kimber). *The logic of failure: Why things go wrong and what we can do to make them right*. New York: Holt.

Dörner, D. & Wearing, A. J. (1995). Complex problem solving: Toward a (computersimulated) theory. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 65-99). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ertmer, P. A., & Quinn, J. (1998). *The ID Casebook: Case studies in instructional design*. Upper Saddle River, NJ: Prentice Hall.

Frensch, P. A. & Funke, J. (1995). Definitions, traditions, and a general framework for understanding complex problem solving. In P. A Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp.3-26). Hillsdale, NJ: Lawrence Erlbaum Associates.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibbons, A. S. (2001). Model-centered instruction. *Journal of Structural Learning and Intelligent Systems*, 14, 511-540.

Goel, V., & Pirolli, P. (1989). Designing within information-processing theory: The design problem space. *AI Magazine, 10*(1), 19-36.

Hayes, J. (1980). *The complete problem solver*. Philadelphia: Franklin Institute Press.

Heppner, P. P. & Krauskopf, C. J. (1987). An information-processing approach to personal problem solving. *The Counseling Psychologist*, 15, 371-447.

Hunt, E. (1991). Some comments on the study of complexity. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jonassen, D. H. (1997). Instructional design model for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology, Research, and Development*, 45(1), 65-95.

MacKinnon, A. J. & Wearing, A. J. (1980). Complexity and decision-making. *Behavioral Science*, 25, 285-296.

MacKinnon, A. J. & Wearing, A. J. (1983). Decision making in dynamic environment. In P. B. Stigum & F. Wenstop (Eds.), *Foundations of utility and risk theory with applications* (pp. 399-422). Dordrecht: The Netherlands: D. Reidel.

Markman, A. B. (1998). *Knowledge representation*. Mahwah, NJ: Erlbaum.

Miller, G. Galanter, E., & Pribram, K. (1960). *Plans and structure of behavior*. New York: Holt.

Milrad, M., Spector, J. M., & Davidsen, P. I. (2002). Model facilitated learning. In S. Naidu (Ed.), *Investigating the effects of information and communications technology: What is working, where and how*. London: Kogan Page.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, 65, 151-166.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ:Prentice-Hall.

Perez, R. S., & Neiderman, E. C. (1992). Modelling the expert developer. In P. R. Chatelier (Ed.), *Advanced technologies applied to training design* (pp. 261-280). New York: Plenum Press.

Perez, R. S., Johnson, J. F., & Emery, C. D. (1995). Instructional design expertise: A cognitive model of design. *Instructional Science, 23*, 321-349.

Putz-Osterloh, W. & Lemme, M. (1987). Knowledge and its intelligent application to problem solving. *German Journal of Psychology*, 11, 286-303.

Rowland, G. (1992). What do instructional designers actually do? An investigation of expert practice. *Performance Improvement Quarterly, 5*(2), 65-86.

Seel, N. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition, and Learning*, 1(1), 59-85.

Sheerer, M. (1963). Problem solving. *Scientific American*, 208, 118-128.

Simon, H. A. (1981). *The sciences of the artifical* (2nd ed.). Cambridge, MA: MIT Press.

Skinner, B. F. (1966). An operant analysis of problem solving. In B. Kleinmuntz (Ed.), *Problem solving: Research, method, and theory* (pp.225-257). New York: Wiley.

Smith, M. U. (1991). A view from biology. In M. U. Smith (Ed.), *Toward a unified theory of problem solving: Views from the content domains* (pp.1-20). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sonenberg, E. A., Lawrence, J. A., & Zelcer, J. (1992). Modeling disturbance management in anesthesia: A preliminary report. *Artificial Intelligence in Medicine*, 1, 447-461.

Spector, J. M., & Anderson, T. M. (Eds.). (2000). *Integrated and holistic perspectives on learning, instruction, and technology: Understanding complexity*. Dordrecht: Kluwer Academic Publishers.

Sternberg, R. J. (1995). Expertise in complex problem solving: Comparison of alternative conceptions. In P. A Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp.295-321).

Hillsdale, NJ: Lawrence Erlbaum Associates.
Wheatley, G. H. (1984). Problem solving in school mathematics. (MEPS Technical Report No. 84.01). West Lafayette, IN: Purdue University, School of Mathematics and Science Center.
Woods, D. T., Crow, C. M., Hoffman, T. W., & Wright, J. D, (1985). Challenges to teaching problem solving skills. *Chem 13 News*, 155, 1-12. Waterloo, Ontario: Canada: University of Waterloo.

Redesigning Graduate Programs in Instructional Design and Technology: Issues, Challenges, & Trends

Deniz Eseryel
Mary Bludnicki
Phil Doughty
Syracuse University

Abstract

The field of instructional design and technology is constantly-expanding and fast-changing. These changes bring about the challenge of redesigning graduate programs that reflect what we value as a field, and where the field is going. In this paper, we present how IDD&E has addressed these challenges during its recent curriculum revision effort. Furthermore, we will provide an overview of some of the most influential graduate programs in US. Reflecting on the trends of the field we, then, discuss how to think about redesigning graduate programs that match up what we value in the field and where the field is going. The paper concludes with operational issues that should be addressed when redesigning programs.

Introduction

Given the new developments affecting our field, and the future of instructional design as predicted from political, economic, social, and technological trends, how does a graduate program that purports to prepare professionals as instructional design and technology researchers and practitioners decide what to teach and how to teach it?

Within the department of Instructional Design, Development, and Evaluation (IDD&E) at Syracuse University, curriculum revision is an ongoing process. The last revision cycle can be said to have started in the Spring semester of 1999. For the past three years, IDD&E has been engaged in an extensive curriculum redesign process in order to be able to reflect current changes in the field in its masters, doctoral, and certificate programs.

To support the curriculum redesign effort at IDD&E, a curriculum planning team, including faculty, graduate students, and alumni was formed. Initial discussions focused upon assessment of the current IDD&E graduate programs. This assessment process involved two components. Informal assessment included reviewing the experiences of students and faculty with existing programs. A more formal assessment was conducted through a systematic analysis of the existing curricula. The curriculum development model used by IDD&E was dynamic in that the analysis, design, development, and evaluation components of a typical systematic curriculum model were integrated throughout every phase in the development model. At the heart of the curriculum redesign effort was the front-end analysis effort (FEA), which included:

- a comprehensive review of graduate and certificate programs in instructional design and technology at 44 competitive institutions in United States;
- the analysis of IDD&E curricula (master's, doctoral, and certificate) against the set of instructional design competencies published by the International Board of Standards for Training, Performance, and Instruction (IBSTPI) (see Richey, Fields, & Foxon, 2000);
- an analysis of market trends through job histories and trends as well as interviews with people in key industries and other organizations;
- input and advice from professional associations such as the Association for Educational Communications & Technology (AECT) and Professors of Instructional Design and Technology (PIDT);
- ongoing discussions, debates, proposals, within IDD&E (faculty, adjuncts, graduate students, alumni, and emeriti).

In this paper, we would like to share our experiences as to how to address the challenges of redesigning graduate programs in instructional design and technology. The following section presents brief results of a study conducted during FEA, which aimed at describing the state-of-the-art of selected graduate programs in instructional design and technology. The next section reviews the trends in the field and provides a view on how to think about designing graduate programs that match up what is valued in the field and where the field is emerging. Nevertheless, there are also operational issues and limitations when it comes to designing or redesigning graduate programs. These include curricular emphasis, contexts to be addressed, sequencing, core course, faculty, schedule, degree offerings, online communication and infrastructure, delivery modes, and so on. These issues are addressed at the concluding section of this paper.

Review Of Graduate Programs In Instructional Design and Technology

The annual reference work Peterson's Graduate Programs in Business, Education, Health, Information Studies, Law & Social Work lists graduate programs in the United States. In the Education division, the section entitled "Administration, Instruction, and Theory" lists "Educational Media/Instructional Technology" programs. The 2000 edition of Peterson's Graduate Programs in Business, Education, Health, Information Studies, Law & Social Work lists two hundred and twenty-three post-baccalaureate degree programs in Educational Media and Instructional Technology. For the purposes of this study, only forty-four graduate programs across United States that offer advanced study (i.e., Master's, Education Specialist, Doctoral) and certificate degrees (i.e., Certificate and Certificate of Advanced Study) in Instructional Design and Technology were reviewed (Bludnicki, 2001).

The following qualifications were used to select the Educational Technology and Instructional Technology programs in the United States:

- Educational institutions that offer Educational Technology/Instructional Technology degrees both at the Master's and Doctoral levels.
- Educational institutions within New York State that offer Educational Technology/Instructional Technology degrees at the Master's level.
- Educational institutions with a four hundred mile radius of Syracuse University that offer Instructional Technology degrees at the Master's level. These educational institutions are located in Massachusetts, New Jersey, Pennsylvania, and Maryland.

Information about the graduate schools was obtained using the following methods:

- Graduate catalogs, brochures
- Educational institutions' Web sites
- Internet Web sites that provided information
- Telephone
- E-mail

The data were collected from the time period the Fall 2000 academic term through the Fall 2001 academic term. The purpose of this descriptive study was to gather information from these forty-four graduate programs in the United States so that it can be evaluated and compared specifically to the Instructional Design, Development and Evaluation (IDD&E) program at Syracuse University.

This descriptive study reviewed the graduate programs, courses, admissions policies, and facilities. It also includes information about students, and faculty. Student information includes career opportunities, student tuition, financial assistance, student enrollment and graduates. Faculty information includes faculty degrees, faculty rank, faculty research, and research funding.

Areas Of Study

The Master's degree programs prepare Instructional Design and Technology students to follow and apply methodologies and models. When students graduate they are competent practitioners. First of all, it was found that nine educational institutions have "Areas of Study." While not defined as a concentration, an "Area of Study." indicates the strengths and direction of an Instructional Technology program. The review of the forty-four graduate programs in Instructional Technology reveals that nine educational institutions are Arizona State University, Bloomsburg University, Brigham Young University, Ithaca College, Purdue University, University of Connecticut, University of South Alabama, University of Southern California, and Wayne State University. Six of the colleges and universities have "Areas of Study' for the master's degree. The master's "Areas of Study" include: design, computers, interactive technologies, training, research, and learning theories. Two of the universities (University of South Alabama and the University of Southern California) have "Areas of Study" at the doctorate level. The doctoral degree prepares students to develop theories and conduct research in appropriate methodologies. The Doctor of Education degree is a professional degree while the Doctor of Philosophy degree is an academic degree.

The doctoral "Areas of Study" include: learning and motivation, instructional systems design, evaluation, research and assessment, technology, and building connections.

Curriculum Concentrations Curriculum Concentrations are offered in 45% (N=20) of the Instructional Technology programs. The twenty colleges and universities that offer curriculum concentration are Arizona State University, Brigham Young University, Duquesne University, Florida State University, Indiana University, Kent State University, New York Institute of Technology, Nova Southeastern University, Ramapo College, Richard Stockton College, Teachers College, Texas Tech University, Towson University, University of Georgia, University of Iowa, University of Maryland, University of New Mexico, University of South Alabama, University of Wyoming, and Wayne State University. At the master's degree level, the curriculum concentrations are offered in these areas: Instructional Design, Multimedia, Business/Training, and Distance Education. At the doctoral level, curriculum concentrations are offered at Brigham Young University in Instructional Design, and Research and Evaluation.

Degree Programs Overall, all of the forty-four educational institutions reviewed have master's degree programs. The Education Specialist degree programs are offered at 27% (N=12) of the universities. Certificate programs are offered at 25% (N=11) of the colleges and universities, while the Certificate of Advanced Study programs are offered at 9% (N=4) universities. However, doctoral degree programs are offered at 70% (N=39) of them. Only 19% (N=6) offer both the Doctor of Philosophy and the Doctor of Education degrees. These educational institutions are: Penn State University, Syracuse University, University of Florida, University of Tennessee, University of Wyoming, and Wayne State University). Next, 26% (N=8) offer the Doctor of Education degree programs. These educational institutions are Boston University, Lehigh University, Northern Illinois University, Nova Southeastern University, Teachers College, Texas Tech University, University of Memphis, and the University of Southern California. Further, only 55% (N=17) offer the Doctor of (Philosophy degree programs. These educational institutions are Arizona State University, Brigham Young University, Florida State University, Georgia State University, Indiana State University, Indiana University (Ed.D. program moratorium), Kent State University, New York University, Ohio University, Purdue University, Texas A&M University, University of Connecticut (Ph.D. program moratorium), University of Georgia, University of Iowa, University of New Mexico (Ed.D. program moratorium), University of South Alabama, and Utah State University.

Core Courses The data reveals that over 66% (N=29) of the forty-four colleges and universities report master's core courses. These universities are Boston University, Brigham Young University, California State University, Duquesne University, Florida State University, Georgia State University, Indiana University, Kent State University, Lehigh University, Penn State University, Purdue University, Richard Stockton College, Rochester Institute of Technology, San Diego State University, Syracuse University, Teachers College, Texas Tech University, Towson University, University of Florida, University of Iowa, University of Maryland, University of Memphis, University of New Mexico, University of South Alabama, University of Southern California, University of Tennessee, University of Wyoming, Utah State University, and Wayne State University.

The master's core course credit-hour requirements reportedly range from six to thirty-six semester-credit hours, or twenty-four to twenty-six quarter-credit hours. Next, the data reveals that 68% (N=21) of the thirty-one universities report doctoral core courses. These universities are Brigham Young University, Florida State University, Georgia State University, Indiana State University, Indiana University, Lehigh University, Ohio University, Penn State University, Syracuse University, Texas A&M University, Texas Tech University, University of Florida, University of Georgia, University of Iowa, University of Memphis, University of New Mexico, University of Southern California, University of Tennessee, University of Wyoming, Utah State University, and Wayne State University.

The doctoral core course credit-hour requirements reportedly range from nine to forty-five semester-credit hours in the Ed.D. Program, and six to forty-seven semester-credit hours in the Ph.D. program. At Ohio University, the Doctor of Philosophy core course credit-hour requirements are twenty to thirty quarter-credit hours.

For the Master core courses, the emphasis seems to be in three main areas: technology tools, instructional design and development, and learning theory. Technology tools include core courses in instructional media, educational television, multimedia, computers, Internet, and distance learning. Instructional design and development includes core courses in instructional systems, instructional technology, analysis and assessment, design, development, evaluation, research, and statistics. Finally, learning theory includes core courses in theories of learning, instruction and curriculum, issues and trends, management, internship or practicum, project or portfolio, and exam.

For the Doctoral core courses, the emphasis seems to be in three main areas: instructional design and development, research and statistics, and foundations of learning. The Instructional Design and Development includes core courses in instructional systems, instructional technology, analysis and assessment, design, development, evaluation, research, and statistics. The foundation of learning includes core courses in foundations of learning, instruction and curriculum, issues and trends, seminar, management and leadership, internship or practicum, and dissertation. Since two-thirds of the forty-four educational institutions studied have master's and doctoral core courses, the core course offerings emphasis and breadth is reflected in the curriculum.

Trends in Instructional Design and Technology

The field of instructional design and technology has dramatically grown and evolved over the past decades. New knowledge has been developed and new directions found. As a result, we, as a field, have been in a constant process of redefining our profession (see Reiser, 2002 for review of definitions of the field). In return, it has been a great challenge to agree upon on the terms used to label the field over the years. These changes brought about another challenge: designing graduate programs that reflect what we value as a field and reflect where the field is going.

Table 1 summarizes the trends in instructional media and instructional design. It is not difficult to see the overlapping between these two areas. The developments in instructional media often affect instructional design and vice versa. The developments in both areas continue to broaden the scope of our field. As these new practices re-shape the instructional design theory and practice, expanding the contexts of application, graduate programs redesign their curriculum to reflect the current trends. As an example, in many ways, the 54-year history of the IDD&E program at Syracuse University provides a reasonable reflection of the changes in the field. This adventure began in 1937 with the establishment of an educational film library in Syracuse followed by the audiovisual service in 1943. Finally, in 1948, the graduate program in Instructional Communications was instituted as one of the first in United States with the first doctoral degree awarded in 1951.

During the 1950s the program emphasized educational filmmaking and instructional television paralleling the audiovisual instructional movement in the field. With developments in the field during 1960s, the emphasis of the program shifted to applied learning theory, programmed instruction, and the systematic design of instruction. Reflecting on these developments, the official title of the program was also changed from Instructional Communications to Instructional Technology. During the 1970s, the program reflected the growing emphasis on systematic design of instruction, motivation, diffusion, evaluation, management, and cost-effectiveness. It was again renamed in 1978 reflecting the changes in the field to be called Instructional Design, Development, and Evaluation, stressing the study of analytical tools and procedures applied in an ever-expanded range of setting. The 1980s marked major changes in the program's curriculum as discussed by Doughty and Durzo (1981) reflecting the impact of cognitive psychology and the use of microcomputers in instructional design. The last curriculum revision cycle in IDD&E brought about many other changes reflecting the recent trends in 2000s (Eseryel & Doughty, 2002).

Similar cases can be made from the history of most other graduate programs. One way to observe how current graduate programs reflect the trends is to simply look at the names of existing graduate programs. Current program names in United States indicate the diversity of graduate study in our field as reflected in Table 1. A recent trend was observed with the inclusion of 'technology' in the program title. Titles emphasizing instructional or educational technology were most frequently used. Some examples are: Educational or Instructional Media and Technology; Educational Communication and Technology; Adult Learning and Technology; Learning and Instructional Technology; Instructional or Educational Technology; Instructional or Educational Media & Technology; Instructional Design and Technology; Curriculum Instruction & Technology; Instructional Psychology & Technology; Instructional Systems Technology; Computer Information Systems; Information Processing Technologies; Instructional Technology & Distance Education; Distance Learning & Media; Instructional Technology Specialist; Technology in Education; Organizational Learning & Instructional Technologies; Multimedia Technologies; Computers in Education; and Instructional Technology Supervisor. Nonetheless, there are a number of programs that limit themselves to the process itself: Instructional Systems; Instructional Systems Development; Instructional Design and Development; Instructional Design, Development, & Evaluation. Still another programs have a different emphasis such as Organizational Communications, Learning, & Design; Organizational Learning & Training; Program Evaluation; Human Resource Development. In most cases, the titles selected indicate important differences, which demonstrate the scope of activities in learning environment development. As the data show programs that include the *technology* in its title tend to emphasize that component of instructional systems more than programs without it.

Table 1. *Trends in instructional design and technology*

	INSTRUCTIONAL MEDIA	**INSTRUCTIONAL DESIGN**
Before 1950	VISUAL INSTRUCTION /VISUAL EDUCATION MOVEMENT (1908-1920s) - Photographs - Magic Lanterns (Lantern Slide Projectors) - Stereopticons (Stereograph viewers) - Motion Picture Projector AUDIOVISUAL INSTRUCTION MOVEMENT (1920s-1930s) - Radio Broadcasting - Sound Recording - Sound Motion-Pictures	VIEW OF TRAINING AS A SYSTEM (1940-1950) - Analysis, design, and evaluation procedures (WWII psychology task groups)
1950s	- Audiovisual Research in Schools - Instructional Television/Educational Broadcasting	- Innovative task analysis methodology (Miller, 1953) - Programmed Instruction (Skinner, 1958) - Popularization of Behavioral Objectives - Taxonomy of Educational Objectives (Bloom et al., 1956)

1960s	- Decline in Educational TV due to low instructional quality	- Criterion-Referenced Testing (Glaser & Klaus, 1962) - The Conditions of Learning: Domain of Learning, Events of Instruction, & Hierarchical Analysis (Gagné, 1965) - Formative & Summative Evaluation (Cronbach, 1963; Scriven, 1967) - Systematic Design of Instructional Materials (Gagné, 1962; Glaser, 1962 & 1965; Silvern, 1964; Barson, 1967; Banathy, 1968, Humerus, 1968) - NEW TERMS: INSTRUCTIONAL DESIGN, SYSTEM DEVELOPMENT, SYSTEMATIC INSTRUCTION, INSTRUCTIONAL SYSTEM
1970s	- The terms 'Educational Technology' & 'Instructional Technology' replaces 'Audiovisual Instruction' - Video-cassettes	- Over 40 ID models developed - Instructional development centers founded in academia - Many graduate program in ID
1980s	- Microcomputers	- Impact of Cognitive Psychology - Computer- Based Instruction - New ID models to accommodate interactive capabilities of computers - Automated Instructional Design
1990s	- '1 computer, in average, for every 9 students in US schools' yet minimal impact of computers in education (drill & practice, and computer-skills teaching only) - Apple computers in K-12 and IBM-compatible computers in universities	- Performance Technology Movement - EPSS - Rapid Prototyping - Constructivism and its Associated Theories (e.g., Activity Theory, Distributed Cognition, Situated Learning, and so on)
2000s	- Internet and related digital technologies - Virtual Learning Environments	- Web-based/ Web-assisted Instruction - Knowledge Management - Performance Engineering - Object-oriented design - Learning Objects - Distributed Learning - Artificial Intelligence - Advancements in Neuroscience

Concentration areas offered at the graduate programs provide a more in-depth look into how the field is reflected in current programs. Reported concentration areas are organized into the following categories of emphasis: process, technology, training, distance education, research & evaluation, K-12, learning, or library media (Table 2). There are, however, three reported concentration areas that do not fit easily into these categories of emphasis. These are: Bilingual Instructional Systems, Science, Communication & Leadership, Research, and Program Evaluation. Not surprisingly, the concentration areas within graduate programs show a wide variety reflecting the broadened scope of the field.

Table 2. *Concentration areas within 44 graduate programs reviewed*

Process	Technology	Training	Distance Education	K-12	Learning	Library Media	Other
Design	Media & Technology	Business	Distance Learning & Education	K-12 Classroom Teach	Learning	Library Media Technology Education	Communication & Leadership
Development	Educational/ Instructional Technology	Corporate	Open & Distance Learning	Classroom Teacher	Teaching & Learning	Library Media Licensure	Science
Implementation	Multimedia in Education	Training & Human Resources Development	Distance Education	Classroom Instruction	Education	School Library Media	ESOL/ Bilingual Instructional Systems
Design, Development, & Evaluation	Instructional Computing	Training Systems		School Instructional Systems	Cognition & Instruction	Information Technology & School Library Media Administration	
Instructional Design & Production	Instructional Technologist	Communications, Training,& Technology		K-12 Education	Organizational Learning		
Performance System Design	Computer Technology Licensure	Performance Improvement & Training					
Instructional Analysis, Design, & Development	Computer Education & Technology						
Implementation & Management	Educational Computing						
Instructional Design & Development	Computer Applications						
Media Design & Production	Multimedia Technologies						
Instructional Development for Training and Education	Instructional Computing						
	Interactive Technologies						

Issues in Redesigning Graduate Programs

It is relatively clear that the field as represented by these forty-four programs does not lack for diversity in emphasis or structure. Common themes and familiar issues can be found throughout the database but exerting some consensus on best practice or exemplary offering or superior scheme is unrealistic. What can be learned from the data is that the field is reasonably focused while also considerably diverse. This applies to both the substance (e.g., goals, context, focus) as well as the form (processes, requirements, schedules) of these programs. In addition, it is plausible that a study similar to this conducted several years from now will reveal considerable changes in most all components. To re-label all of these programs as the 'Department of Changing Minds' may be an accurate description of what is occurring. Whose minds and what changes will be interesting to consider over time.

The following list of issues to be considered reflects perspectives of individuals new to the field, faculty seeking guidance on course development, academic programs comparing themselves to others, and administrators deciding upon necessary and sufficient resources to generate such programs:

- Focus or Emphasis
  - Certificate (local or state registration)
  - Degrees (BS/MA/MS/CAS/Ph.D./Ed.D.)
  - Concentrations and Specializations
  - Theory, Practice, Inquiry, Field-based
- Resources
  - Faculty
  - Staff
  - Facilities
  - Web Support
  - Financial Support

  - Student Population
- Structure
  - Admission Requirements
  - Credit Requirements
  - Sequence
  - Core Requirements
  - Assessment Processes
  - Exit Requirements
  - Residence
  - Full/Part Time
  - Day/Evening Offerings
  - Summer Programs
- Contexts
  - K-12
  - Higher Education
  - Business and Industry
  - Government
  - Not-for-Profit
  - Contract R&D
  - Internet
- Process
  - Face-to-Face
  - Blended
  - On-line
  - Schedule
  - Practicum/Internship
- Roles
  - Analyst
  - Designer
  - Developer
  - Instructor
  - Evaluator
  - Manager
  - Researcher

Each of the above six categories of issues can be considered in isolation but in fact, the mix between and among them in the forty-four programs is extensive. Unfortunately, the page-limitation of this document does not allow us to further comment on these issues in more detail. However, as a concluding comment, we emphasize the importance of addressing these issues during curriculum redesign and having extended conversations with other departments about relationships and connections throughout the list.

References

Banathy, B. H. (1968). *Instructional systems*. Belmont, CA: Fearon.

Barson, J. (1967). *Instructional systems development. A demonstration and evaluation project: Final report*. East Lansing, MI: Michigan State University. (ERIC Document Reproduction Service No. ED 020 673)

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *The taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive Domain*. New York: David McKay.

Bludnicki, M. (2001). *A study of instructional technology graduate programs at selected educational institutions in the United States*. Unpublished manuscript, Syracuse, NY.

Cronbach, L. J. (1963). Course improvement through evaluation. *Teachers' College Record*, 64,672-683.

Doughty, P. & Durzo, J. (1981). Redesign of a graduate program in instructional design, development, and evaluation (IDDE). *Journal of Instructional Development*, 4(3), 14-20.

Eseryel, D & Doughty, P. (2002). *Does our curriculum match up? Redesigning a graduate program in instructional design and technology*. Paper presented at the 2002 Professors of Instructional Design and Technology Annual Meeting, Smith Lake, VA.

Gagné, R. M. (1962). (Ed.), *Psychological principles in system development*. New York: Holt, Rinehart, & Winston.

Gagné, R. M. (1965). *The conditions of learning* (1st ed.). New York: Holt, Rinehart, & Winston.

Glaser, R. (1962). Psychology and instructional technology. In R. Glaser (Ed.), *Training research and* *education*. Pittsburgh: University of Pittsburgh Press.

Glaser, R. (1965). Toward a behavioral science base for instructional design. In R. Glaser (Ed.), *Teaching machines and programmed learning: Vol. II. Data and directions*. Washington, DC: National Education Association.

Glaser, R. & Klaus, D. J. (1962). Proficiency measurement: Assessing human performance. In R. M. Gagné (Ed.), *Psychological principles in system development*. New York: Holt, Rinehart, & Winston.

Hamerus, D. (1968). *The systems approach to instructional development: The contribution of behavioral science to instructional technology*. Monmouth, OR: Oregon State System of Higher Education, Teaching Research Division.

Mager, R. F. (1962). *Preparing objectives for programmed instruction*. Belmont, CA: Fearon.

Miller, R. B., (1953). *A method for man-machine task analysis* (Tech. Rep. No. 53-137). Wright-Patterson Air Force Base, OH: Wright Air Development Center.

Reiser, R. A. (2002). What field did you say you were in? Defining and naming our field. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (pp. 5-15), NJ: Merrill Prentice Hall.

Richey, R., Fields, D. C., & Foxon, M. (2000). *Instructional design competencies: The standards* (3rd ed.). Syracuse, NY: ERIC Clearinghouse on Information and Technology in cooperation with the International Board of Standards for Training, Performance, and Instruction (IBSTPI).

Scriven, M. (1967). The methodology of evaluation. In *Perspectives of curriculum evaluation* (American Educational Research Association Monograph Series on Curriculum Evaluation, No.1). Chicago: Rand McNally.

Silvern, L. C. (1964). *Designing instructional systems*. Los Angeles: Education and Training Consultants.

Skinner, B. F. (1958). Teaching machines. *Science*, 128, 969-977.

The Benefits of an Efficiency Metric in the Evaluation of Task Performance

David F. Feldon
University of Southern California

Abstract

Most current measures of performance in training contexts rely on composite or binary scores derived from sub-goals achieved by the learner. However such techniques provide little information about the learner's skill acquisition stage or progress towards expertise attained during the task itself. The application of a three stage model of skill acquisition is examined (Anderson, 1982, 1995; Fitts & Posner, 1967), and a combination of three efficiency metrics is proposed to elucidate learners' progress towards automated expert performance. Implications for computer adaptive testing and just-in-time learner support are discussed.

Objectives

At a fundamental level, many current measures of task performance in training simulations simply calculate the sum of the number of sub-goals successfully attained by the learner in pursuit of the primary goal objective. This approach, while effective for monitoring a user's ability to complete a task, is limited in its ability to provide richer information about the quality and stability of the skill that the user is acquiring. Cognitive models of skill acquisition suggest that as learners progress from novice to expert, procedural knowledge becomes automated, conserving limited attentional resources. However, in the process of automating, the procedural skills become highly resistant to modification (Anderson, 1983). Such improvements in cognitive or mental efficiency (Cobb, 1997; Paas & van Merrienboer, 1993) indicate high levels of proficiency that are not observable using mechanisms that only evaluate tasks and sub-tasks as complete or incomplete, because there is no means of capturing the amount of mental effort required by the user to complete the task. The purpose of this paper is to provide theoretical support for an additional assessment method that evaluates learner progress in training simulations using three simultaneous measures of task efficiency that, when evaluated in parallel, can support highly adaptive advanced simulation training based on cognitive modeling and provide a stronger assessment of learning outcomes.

Theoretical Framework

According to Cannon-Bowers and Salas (1997), four basic properties should guide the selection of performance indicators in training environments: the ability to assess multilevel task components in accordance with the demands of the training; the ability to assess both the training task's outcome and process; the ability to provide descriptive, evaluative, and diagnostic data on the performance; and the ability to inform decisions regarding training advancement and remediation. While any number of measurement systems have been developed that meet these criteria (e.g. TARGET; Fowlkes, Lane, Salas, Franz, & Oser, 1994; SALIANT; Muñiz, Stout, Bowers, & Salas, 1998; UPAS; Meliza & Tan, 1996), each produces a metric that is fundamentally binary in nature. That is, specific goals and sub-goals can only be reported as completed or failed. As such, the information that they can generate for a particular user is limited to individual event scores or combinational percentages. Although such measures are useful to a certain extent, aggregate linear composite measures are quite limited in the depth and richness of the data they provide (Gipps, 1994).

As training simulations are designed to facilitate users' progress toward expertise, it follows that an optimal performance metric would be able to inform evaluators of an individual's fluency in the completion of a given task as well as indicators of success and failure. Such data is especially valuable in the context of intelligent tutoring and adaptive systems, when the differential effects of instruction relative to learners' skill levels are considered. For example, the use of worked examples for novices is significantly more effective than direct instruction (Trafton & Reiser, 1993) or unstructured exploration in simulated environments (Kalyuga, Chandler, & Sweller, 2001; Rieber & Parmley, 1995). On the other hand, more advanced learners learn more effectively from instructional environments with fewer explicit instructional supports (Kalyuga, Chandler, Tuovinen, & Sweller, 2001; Mayer, 2001).

The stage model of skill acquisition delineates three identifiable levels of skill mastery: the cognitive stage, associative stage, and autonomous stage (Anderson, 1982, 1995; Fitts & Posner, 1967). The first stage (cognitive) represents the learner's dependence upon explicit declarative knowledge of prescribed steps to complete a defined task. Such declarative knowledge need not be verbalized instructions, per se. Rather, learners' recall of examples has been found to scaffold the development of new skills. Further, use of these examples is evident, even if the examples themselves are not committed to long term memory (Anderson & Fincham, 1994). Thus, even when examples have only been recently introduced, they can provide a conscious model for the successful resolution of problems in early stages of learning. Anderson, Fincham, and Douglass (1997), for example, demonstrate that subjects trained on a reasoning task using only examples without having the benefit of stated rules or principles begin to successfully solve problems by analogically extending the training examples provided. They further note many studies by other researchers which have reported that despite being taught rules and procedures for completing tasks, learners typically make explicit references to available examples, even when adequate procedures for solving the problem were provided (e.g. Chi, Bassock, Lewis, Reimann, & Glaser, 1989; Novick & Holyoak, 1991; Ross & Kennedy, 1990).

In the second stage (associative), the declarative knowledge of the procedure transitions to a proceduralized state, wherein conceptual understanding develops and execution is relatively accurate, though vulnerability to changes in the environment that necessitate adjustments to the order of the steps to task completion remain. Anderson and Fincham (1994) note that "there seems to be a gradual shift from example-based processing to rule-based processing" (p. 1338) as subjects practiced programming skills over repeated trials. This gradual transition also enhances learners' ability to transfer new skills by generalizing the appropriate procedural mechanisms to a point where they will be effective with problems for which the learner has not had specific instructions or models presented.

At the final (autonomous) stage of expertise, procedural knowledge is automated, such that task execution is both adaptable to changing conditions and rapidly performed with minimal conscious mental effort. It is at this stage that flexibility is at once greatest and most restricted. Skill transfer capabilities are highly adaptable to complex, ill-structured, and novel situations, because the automated procedure occupies minimal space in working memory, which allows much more mental effort to be spent attending to relevant new details (Jonassen, 2000) . However, the procedure itself becomes more ingrained and extremely difficult to change to the extent that it can manifest without conscious activation (Aarts & Dijksterhuis, 2000). As such, once a skill has been automated, it no longer operates in such a way that it is available to conscious monitoring, and it tends to run to completion without interruption, further limiting the ability to modify performance (Bargh & Ferguson, 2000; Wheatley & Wegner, 2001).

Analysis of the Literature

VanLehn (1996) observed that most assessments of cognitive skill acquisition do not evaluate performance during the cognitive stage. As mentioned above, this stage entails a great deal of declarative knowledge involvement in the development of procedures, which results in a slow initial rate of performance that increases with practice. The underlying organization of the process has not become stabilized as procedural knowledge, and rapid performance is hindered due to the necessary dependence on conscious decision making, monitoring, and reprocessing (Segalowitz, 2000). In their study of complex skill acquisition, Anderson and Fincham (1994) demonstrated that both speed and accuracy increase as subjects become less dependent on declarative representations of the examples and instructions provided. Their findings also indicated that procedural speed increases very sharply and then stabilizes as the production (i.e. procedure) is practiced without reference to declarative representations of the method for completing the task. Rephrased in the terminology of the three-stage model, the learners progress to the associative stage of skill acquisition.

In an attempt to examine this type of transition in finer detail, Segalowitz and Segalowitz (1993) found through their study of second language word recognition tasks that a reliable measure of performance at the cognitive stage and, as proficiency increases, an indicator of a learner's transition to the associative stage is the relative variability of response time in performance, computed by dividing each subject's standard deviations of response time by the individual's mean response time. As speed of performance accelerates in the first stage, the standard deviation and the mean time decrease proportionately, but the relative variability remains constant. When the second stage is reached, however, the process no longer relies on declarative knowledge, so without additional "mental instruction" to modify the procedure, the standard deviation of response time remains constant while the mean response time continues to decrease, resulting in a diminishing measure of relative variability.

Another important element of complex skill assessment that is often neglected is a measure of task efficiency (Feldon, 1998; O'Neil, 2003). In realistic situations, there is often added cost associated with a procedure that, while effective, involves unnecessary steps. Skilled workers paid at an hourly rate cost more to employ if they utilize inefficient procedures to complete their tasks, because the work takes longer to complete. In higher stakes endeavors, such as surgery or combat, the time spent on unnecessary steps can significantly increase the risk of complication or injury. As Carr (1991) explains, "the [problematic] process will be inefficient. That is, it will use too many resources for the amount of output it produces… The right output may be produced, but it will be too expensive or too late for its intended purpose" (p. 44).

Measures of efficiency for task completion have been used, for example, to evaluate the quality of computer code generated by students in programming classes. Such variables as number of lines of code and numbers of necessary compilations for error-free use have been found to be valid indicators of student skill (Hung, Kwok, & Chan, 1993). In a related approach, Hong and Liu (2003) utilized a direct step-counting measure of procedural efficiency to differentiate expert and novice performance for their study of expertise in abstract problem-solving games. For each mouse click or keystroke that was made in the completion of a game level, a counter increased by one. Their classification of experts was based on the generation of a significantly lower number necessary to complete a given level than novices playing the same level. Such step-counts can be transformed into measures of efficiency by dividing the total number of user steps by the minimum number of steps required to complete the task.

Because the key distinction between the associative (second) and autonomous (third) stages of skill acquisition is a decrease in mental effort for the completion of a given task (i.e. an increase in mental efficiency), rather than a reliable difference in the accuracy of performance, a composite score of the nature described earlier is also limited in its ability to definitively distinguish between a stage two and stage three learner. As noted by Camp, Paas, Rikers, and van Merrienboer (2001), "when only performance is used as a selection parameter, no difference is made between people who perform well and indicate a high mental effort and people who also perform well, but indicate a low mental effort" (pp. 579-580). Mental effort, considered an index of cognitive load (Paas, 1992), is minimized due to the completed formation of schemas, which organize and categorize new information, and the automatic procedures that have developed for processing it (Chi, Glaser, & Rees, 1982). In Camp, et al.'s (2001) study of computer adaptive problem selection for air traffic control training, the authors found that utilizing a mental efficiency metric representing the discrepancy between standardized mental effort and performance scores on a given problem to govern the automated selection of training problems led to the most significant gains in complexity of problems successfully solved.

Conclusions

As Kelley, Bosman, Charness, and Mottram (1998) observed in their examination of proficiency measurement techniques, "In some contexts, multiple evaluation techniques can and should be used" (p. 35). When considered independently, each of the skill assessment techniques described above provides only limited information about the state of the learner's expertise when used in isolation. For instance, speed of performance can indicate transition from the cognitive stage, but provide little insight into the refinement of the procedure or the degree of automaticity. Likewise, efficiency of task completion in terms of minimizing the number of steps in a task cannot inform an evaluator of the skill that is encoded independently of declarative support, and efficiency of mental effort cannot reveal the efficacy of the procedure. However, when considered with the completion pattern of tasks and subtasks, gathering and analyzing the three efficiency metrics collectively in real time, a comprehensive picture emerges that not only indicates what stage of skill acquisition the learner is currently moving through, but also can be extremely useful for computer-adaptive problem selection and just-in-time supportive instruction to enhance learner performance along each of the three axes: speed, task efficiency, and mental effort (automaticity).

Implications

This system of assessment allows for highly sophisticated intervention strategies to be integrated into intelligent tutoring systems. By monitoring task completion efficiency, for example, it is possible to remediate inefficient procedures before they become automated and resistant to change. For example, if a learner's level of task completion efficiency had not yet reached acceptable levels and an initial increase in the level of mental efficiency was detected (i.e. the inefficient procedure was beginning to automate), just-in-time instruction could be delivered to improve the task completion efficiency before additional practice problems that led to further automatization of the skill were offered.

Further, subtask performances could also be analyzed independently within the context of completion for the higher level task. Analyzing all three efficiency types for each sub-component of a task would allow for a deeper understanding of the procedural acquisition stage of each discrete skill. Specifically, failure to complete a high level task satisfactorily could be understood in terms of the time, steps, and mental effort dedicated to each sub-component, rather than losing such detail through obfuscation by traditional linear composite and binary measures. Developing a model of the learner at this level of specificity could inform further, highly specific training that was carefully tailored to address only the particular areas of weakness without unnecessarily dedicating instructional resources to those elements executed at an acceptable level of performance.

References

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. Journal of Personality and Social Psychology, 78(1), 53-63.

Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Anderson, J. R. (1995). Cognitive psychology and its implications (4th ed.). New York: Freeman.

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(6), 1322-1340.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(4),

932-945.
Bargh, J.A. & Ferguson, M,J. (2000). Beyond Behaviorism: On the Automaticity of Higher Mental Processes. Psychological Bulletin, 126(6), 925-945.
Camp, G., Paas, F., Rikers, R., & van Merrienboer, J. (2001). Dynamic problem selection in air traffic control training: a comparison between performance, mental effort and mental efficiency. Computers in Human Behavior, 17, 575-595.
Cannon-Bowers, J. A., & Salas, E. (1997). A framework for developing team performance measures in training. In M. T. Brannick, E. Salas, & C. Prince (Eds.), Team performance assessment and measurement: Theory, methods, and applications (pp. 45-62). Mahwah, NJ: Lawrence Erlbaum Associates.
Carr, C. (1991). How performance happens (and how to help it better), Part 2. Performance and Instruction, 30(1), 42-46.
Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), Advances in the psychology of human intelligence. Hillsdale, NJ: Erlbaum.
Chi, M., Bassock, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.
Cobb, T. (1997). Cognitive efficiency: Toward a revised theory of media. Educational Technology Research and Development, 45, 21-35.
Feldon, D. F. (1998). Efficiency in task completion as a measure of conceptual understanding in RIDES student modeling. Technical Paper. Redondo Beach, California: University of Southern California, Behavioral Technology Laboratories, 1998.
Fitts, P. M., & Posner, M. I. (1967). Human Performance. Belmont, CA: Brooks-Cole.
Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETs methodology. Military Psychology, 11, 189-215.
Gipps, C. V. (1994). Beyond testing: Towards a theory of educational assessment. Washington, DC: The Falmer Press.
Hong, J., & Liu, M. (2003). A study on thinking strategy between experts and novices of computer games. Computers in Human Behavior, 19, 245-258.
Hung, S., Kwok, L., & Chan, R. (1993). Automatic programming assessment. Computers and Education, 20(2), 183-190.
Jonassen, D. H. (2000). Toward a meta-theory of problem solving. Educational Technology: Research and Development, 48(4), 63-85.
Kalyuga, S., Chandler, P., & Sweller, J. (2001). Learner experience and efficiency of instructional guidance. Educational Psychology, 21, 5-23.
Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. Journal of Educational Psychology, 93, 579-588.
Kelley, C. L., Bosman, E. A., Charness, N. H., & Mottram, M. A. (1998). Reliability assessment of a user proficiency measurement technique. International Journal of Human-Computer Interaction, 10(1), 33-49.
Mayer, R. (2001). Multimedia Learning. New York: Cambridge University Press.
Meliza, L. L., & Tan, S. C. (1996). SIMNET unit performance assessment system (UPAS) version 2.5 user's guide (ARI Research Product 96-05). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
Muñiz, E. J., Stout, R. J., Bowers, C. A., & Salas, E. (1998). A methodology for measuring team situational awareness: Situational awareness linked indicators adapted to novel tasks (SALIENT). Proceedings of the First Human Factors & Medicine Panel Symposium on Collaborative Crew Performance in Complex Operational Systems. Neuilly-sur-Seine, France: Research and Technology Agency.
Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 398-416.
O'Neil, Jr., H. F. (2003). Computer based assessment of learning. Presented at the Annual Meeting of the American Educational Research Association, Chicago, IL: April 24, 2003.
Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. Journal of Educational Psychology, 84(4), 429-434.
Paas, F., & van Merrienboer, J. (1993). The efficiency of instructional conditions: an approach to combine mental effort and performance measures. Human Factors, 35(4), 737-743.
Rieber, L. P., & Parmley, M. W. (1995). To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science. Journal of Educational Computing Research, 13, 359-374.
Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 42-55.
Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance (Chapter 11). In H. Riggenbach (Ed.), Perspectives on fluency (pp. 200-219). Ann Arbor, MI: University of Michigan Press.

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and differentiation of speed-up from automatization effects: Evidence from second language word recognition. Applied Psycholinguistics, 14, 369-385.

Trafton, J. G, & Reiser, B. J. (1993). The contribution of studying examples and solving problems to skill acquisition. Proceedings of the 15th Annual Conference of the Cognitive Science Society (pp. 1017-1022). Hillsdale, NJ: Lawrence Erlbaum Associates.

VanLehn, K. Cognitive skill acquisition. Annual Review of Psychology, 47, 513-539.

Wheatley, T., & Wegner, D. M. (2001). Automaticity of action, Psychology of. International Encyclopedia of the Social and Behavioral Sciences, 991-993.

Varying Text Context in Computer-Based Training for Adults in the Workplace

Kathryn M. Fischer
Wilhelmina C. Savenye
Howard J. Sullivan
Arizona State University

Abstract

This research study compared achievement, program time, and attitudes of 78 employees at a large university who completed a computer-based training program for an on-line purchasing and financial system. Subjects were randomly assigned to three treatments: 1) a generic or non-referenced (NR) version that is typical of training environments, 2) an other-referenced (OR) version that is a story-themed version presented through a fictional character who works in the Music department, and 3) a self-referenced (OR) version of the same story setting with the study participant as the character attending training. Subjects in the OR treatment group achieved significantly higher scores on the posttest than subjects in the NR treatment group. Subjects in the SR treatment group spent significantly less time in the program that subjects in both the NR and the OR treatment group. Results suggest that using stories or scenarios provides a way to improve achievement for adults using CBT.

The use of a story or scenarios as an instructional device is well accepted. Brown, Collins, & Duguid (1989) suggested using stories to create an appropriate setting for situating learning in a realistic work environment. Instructional designers and teachers often embed new content in a scenario so that it can be practiced and learned in a more realistic applied setting (Smith & Birney, 1997; McLellan, 1993).

Story lines and scenarios have been used in computer-based training (CBT) in a variety of training and work environments. Tam, Wedd, and McKercher (1997) employed narrative accounts of realistic characters in a computer-based instructional program to teach the concepts and components of a rural accounting system. Gathany & Stehr-Green (1994) developed a CBT using a story of a real life situation and appealing fictional characters to teach medical personnel the principles of conducting epidemiological outbreak investigations. In a similar manner, realistic scenarios and artifacts were incorporated into a CBT course to teach weather forecasters the complex task of interpreting Doppler radar images (Casey, 1996).

An important issue regarding the use of stories and scenarios in instruction or training relates to point of view presented in the learning materials. The point of view in stories commonly is that of others, namely the characters in the story. Don (1990) recommends using a story with characters that have a specific point of view as a way to meld multimedia content and interfaces into an interrelated and cohesive instructional program. In contrast, an alternate point of view, that of the learner or trainee, also is often used in scenarios in which new learning content or problems are embedded. It is not clear which, if either, of these points of view is a more effective and appealing one for use in training and instructional programs that teach learners to perform complex tasks or use complex systems.

Many studies have compared self-referencing (SR) to other-referencing (OR) or to a more abstract reference, e.g., semantic processing. Generally, the terms self-referencing refers to the study subject in the second person, usually as you. Other-referencing refers to persons other than the study subject, using the third person and referring to them by name (John) or a descriptor, such as your friend or your mother. Other-referencing studies have made distinctions between very familiar or intimate others and less familiar others because events and knowledge of intimate others is often interconnected with the self, and thus exhibits similar characteristics to self-referencing.

Free and cued recall studies from word lists with trait adjectives and nouns that compared self-referencing to semantic processing have resulted in better performance under self-referencing conditions (Bower & Gilligan, 1979; Keenan & Baillet, 1980; Rogers, Kuiper, & Kirker, 1977). Not all studies have favored SR. Some studies have indicated that other types of processing promote recall as well as or better than self-referencing (Bellezza & Hoyt, 1992; Keenan, Golding, & Brown, 1992; Klein & Kihlstrom, 1986; Maki & McCaul, 1985).

Lord (1980) obtained results that failed to support the self-referencing advantage in studies that compared the recall of information associated with the subjects (SR), a very familiar other (OR), such as the subject's father; a less familiar person (OR) and an inanimate object. Subjects were asked to imagine a picture of an object represented by a given word and an interesting picture of a person (self, familiar other, less familiar other, or inanimate object) interacting with the object for the given word. In contrast, Brown, Keenan, and Potts (1986) presented results from a series of imagery studies that contradicted Lords' findings (1980). In a follow-up imagery study, Lord (1987) obtained results that again supported the other-referencing advantage.

Because of reported inconsistencies in findings for superior recall under SR conditions versus findings of equal or better recall under non self-referencing, Symons and Johnson (1997) conducted a meta-analysis of 129 such word studies, 126 of which used adult subjects. They reported being unable to counter the findings for OR in Lord's study (1980). They also concluded that significant positive differences in achievement, transfer, and depth of processing were attributable to the self-referencing effect.

A study by Reeder, McCormick, and Esselman (1987) differed from earlier self-referencing and other-referencing studies. Instead of using word lists (trait adjectives and nouns) to compare recall, Reeder et al. used 100-word reading passages under self-reference, other-reference, linguistic, and control conditions. Results of the first of two experiments indicated significantly better recall for the SR and OR conditions over the linguistic and control conditions. Results of the second experiment indicated significantly better recall for the SR over the other three conditions.

There are potential advantages of both other-referenced and self-referenced training and instructional programs. One advantage of the other-referenced training program is that learners may find it more realistic to visualize a character other than themselves as the integrator of new content into a particular setting. They may view this character as a more experienced person from whom they can learn and may be motivated by the positive attitudes and successes of the character.

Potential advantages can also be cited for a self-referenced environment. One is that SR establishes a context in which the story or scenarios are referenced more closely to the learner. Self-referenced training or instructional programs generally use the term you in referring to the learner. This self-orientation is more consistent with personalization of instruction, which has been found to be more effective than non-personalized instruction in several studies (Anand & Ross, 1987; Davis-Dorsey, Ross, & Morrison, 1991; Ku & Sullivan, 2002; Lopez & Sullivan, 1992). Studies that have involved self-referencing through the use of personalization treatment have generally yielded positive attitudes toward the personalization treatment (Cordova & Leper, 1996; Dwyer, 1996; Ku & Sullivan, 2000, 2002). A problem with employing personalization, especially in longer or more complex instructional programs, is that programs that are complex or that target large numbers of learners may not be well-suited to personalization.

Greco (1996) reported that the learning effect of the use of stories or story-like instruction as an instructional device has seldom been measured, even though she views stories as powerful teaching tools both in schools and in the workplace. For example, attitudes are measured much more often than achievement in workplace and training settings. Results of a 1996 study of corporate training programs indicated that only four percent of the responding organizations measured learning results, while 89 percent indicated that they measured attitudes (American Society for Training and Development, 1996).

No studies were found that combined the elements of CBT in the workplace, with non-story and story versions using both an other-referencing context and a self-referencing context. Most studies on SR and OR were not computer-based and have had treatment times of only a few minutes in which subjects were expected to recall words that had little or no further application in their everyday lives.

In contrast, the present study was designed to investigate the effects of self-referenced, other-referenced, and non-referenced (generic) computer-based training on the achievement and attitudes of adult learners being trained to perform complex tasks relevant to their workplace setting in a major university. The study differed from most research on self-referencing and other-referencing in that the learning tasks were much more complex, the treatments consisted of training on important everyday workplace tasks, the average time of the training/treatment mode was more than two hours, and a typical non-referenced version was included with self-referenced and other-referenced treatments.

For this study, the other-referenced (OR) version was developed using a story set in the Music department at the university where this study was conducted. The main character, Bea Sharpe, is a newly-hired pleasant, professional employee, who appears eager and motivated to learn her new job. Study participants experienced the training from the point of view of seeing how another person, Bea, is learning to perform authentic tasks and activities related to using the *Advantage* Financial system for ordering goods and services for the Music department.

The self-referenced (SR) version used a story set in the same university Music department. The study participants were asked to imagine themselves as new employees in the Music department who are learning to perform the same authentic tasks and activities related to using the *Advantage* Financial system for ordering goods and services for the Music department, as Bea did in the OR version. In the SR version, the term *Bea* was replaced by the word *you* and the learners were addressed as if in a face-to-face conversation with an unseen instructor.

The non-referenced (NR) version employed neither a story nor a Music department setting. The non-referenced version used the same learning tasks and number and types of examples, but without the specific references to the characters or Music department examples. In one learning scenario that was related to conference registration fees, there were neither references to particular faculty members, nor to the name or location of the conference, as were used in the OR and SR story versions.

The following research questions were investigated in the study.

1. Is there a difference in learner posttest achievement in self-referenced, other-referenced, and non-referenced training versions?
2. Is there a difference in learner achievement on en route quizzes in self-referenced, other-referenced, and non-referenced training versions?
3. Is there a difference in learner attitudes in self-referenced, other-referenced, and non-referenced training versions?
4. Is there a difference in program completion times in self-referenced, other-referenced, and non-referenced training versions?

It was expected that the SR and OR versions would yield greater achievement and more positive attitudes that the more generic NR version. Evidence generally indicates that more story-like settings produce greater engagement with the learning task and higher learner appeal. There were no clear advance expectations with regard to the relative effects of the SR and OR versions.

This study was conducted in a training setting with subjects who would normally participate in the training. The facilities, equipment, and facilitators were those normally used for this computer-based training. The training was conducted over a 6-month period. Classes were held one or two times per month, depending on the number of participants requesting the class. In total, 12 classes were conducted with a range of 6 to 11subjects per class. Each class session was scheduled for a 4-hour period.

Method

Subjects

Subjects for this study were 78 employees of a large southwestern public university who signed up for training on *Advantage*, the university's on-line purchasing and financial system. The participants included 16 males and 62 females, ages 20-59. Thirty-eight participants (49%) were under 40 years of age, and 40 (51%) were age 40 or above. Thirty-nine participants had a high school diploma or some college and 39 had a bachelor's or advanced degree. Thirty-seven participants (47%) were secretaries or stenographers, 10% were finance clerks or finance specialists, and 6% were faculty or scientists.

Materials

Three parallel versions of materials for this study (non-referenced, other-referenced, and self-referenced) were adapted a half-day computer-based training course entitled Introduction to Advantage, Part A, which is the first in a series of training courses in the use of the Advantage on-line purchasing and financial system. It is offered to faculty and staff who work at the university. Participants must complete this course and others to receive the necessary security access to the Advantage financial system. Financial-system users generally use the on-line Advantage system to order goods and services from on-campus and off-campus suppliers, to look up information related to the status of orders, and to monitor and maintain budget information on various types of accounts. Since its inception in 1998, this course had been through several revisions based on extensive field testing and formative evaluation (Fischer & Ku, 2000; Fischer, Savenye, & Sullivan, 2002).

The unit responsible for Advantage training at the university used Asymetrix Toolbook Instructor II and Microsoft Access to develop the CBT for this study. The CBT incorporates the instructional modules, the posttest, and a custom-developed course management system that is used to manage learners' access to the CBT, to track individual learner progress and achievement, and to collect data for later analyses. The CBT includes 316 base screens that are identified by type, including instruction, review, practice with multiple-try feedback, scored quizzes, and glossary. Thirty-six of these screens contain examples and additional information that are available on the base screen by moving the mouse to a check mark box, which causes the additional information to appear at the bottom of the screen.

The learner begins the program with a Quick Tour that provides instruction for the essential learner-control elements used in navigation and special features, such as hotwords and the My Progress tracking feature that is part of the course management system. After the Quick Tour, the learner proceeds through each of five content modules in order. These modules are titled Overview, Acquiring Goods, Account Codes, Suspense File, and Accessing Advantage. Each of these modules includes instruction, review, practice, and a quiz in either the non-referenced, other-referenced, or self-referenced mode.

The other-referencing version is an account of a new employee, Bea Sharpe, who needs to use *Advantage* to complete tasks related to ordering goods and services for her supervisor in the Music department. The narrator and graphics present the story to the learners so that they appear to be viewing what Bea does as the content is presented. The name Bea is used extensively in the OR version throughout the instruction, review, practice, quiz, and feedback segments.

The self-referencing presents the instruction is presented from the second-person point of view. The self-referencing version uses the same Music department setting and story line, but the learner (referred to as *you*) is involved in the story, replacing Bea as the character performing the learning tasks. Other slight wording

changes were made between the SR and OR versions to maintain appropriate syntax.
The non-referenced version does not use a story or refer to characters or participants explicitly or by name, while the other-referencing and self-referencing versions have a story line with a main character and other supporting characters. The non-referenced version presents the content in an impersonal or generic manner that is typical of many training programs.

Examples of two parallel instruction screens from each of the three versions are shown in Figures 1-6 below.

Review screens present a summary of the instructional content for each unit using a notebook page for the background. Practice screens follow the review screens for each instructional unit. Practice screens have multiple-choice statements with four selections for each practice item, except for a unit in the Account codes module, which has only three possible choices. Each of the answer selections has feedback adapted to the answer choice selected by the participant. The feedback is varied for each of the versions in a manner similar to instruction and review screens.

Four screens that are not in the non-referenced version were added to the other-referencing and self-referencing versions. In the other-referencing version, these screens are used to introduce the main character, Bea, and her workplace department. In the self-referencing version, these screens situated the study participant in the Music department setting.

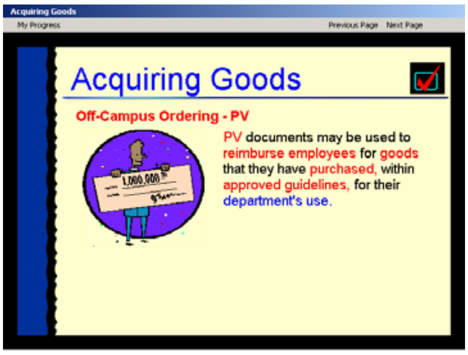


Figure 1. Instruction screen example in the non-referenced version.

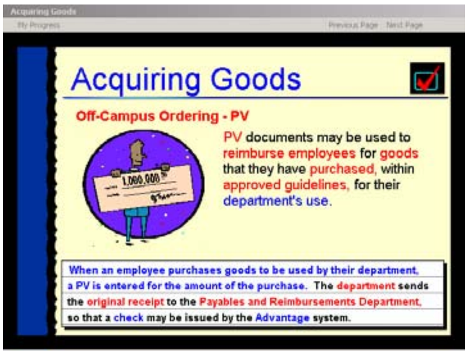


Figure 2. Additional information in the non-referenced version.

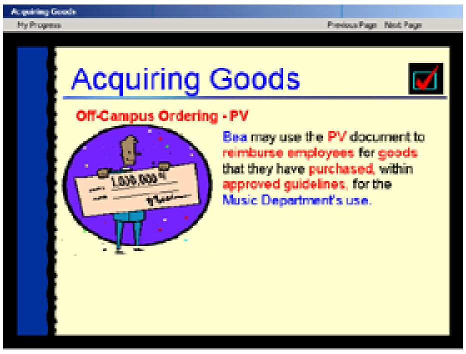


Figure 3. Instruction screen example in the other-referenced version.

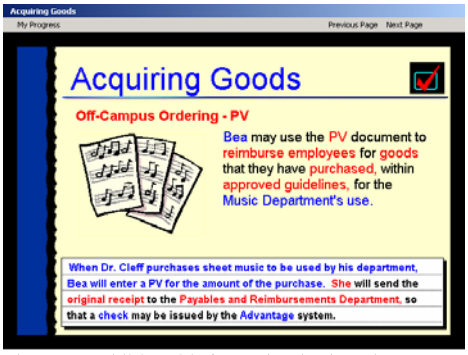


Figure 4. Additional information in the other-referenced version.

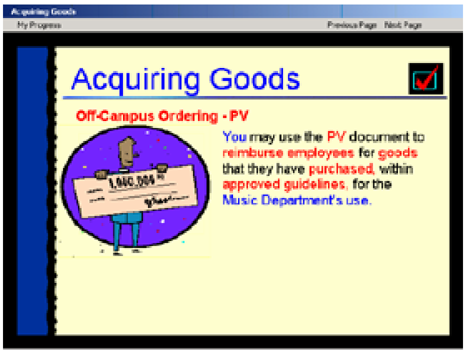


Figure 5. Instruction screen example in the self-referenced version.

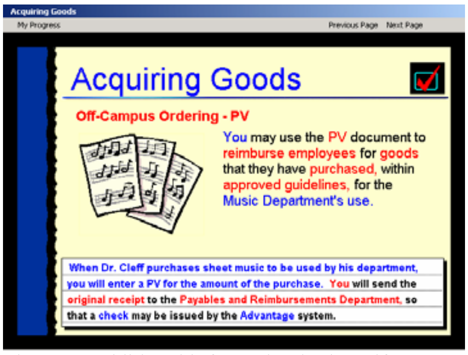


Figure 6. Additional information in the self-referenced version.

Procedures

As is typical in the organization, participants requested training by contacting the *Advantage* Helpline. Participant data such as name, phone number, work department, employee identification number, etc., were collected in a training database for scheduling purposes. Classes were scheduled when there were enough potential learners to conduct a class of up to 11 participants. Each class was limited to a maximum of eleven participants based on the equipment available in the training facility. Subjects were randomly assigned to the three treatments within each class session, after they were scheduled for training.

At the beginning of each training session, the facilitator greeted learners as they entered the training facility. After the facilitator's welcome and introductions, the participants introduced themselves. The facilitator discussed the *Introduction to Advantage* course, pointing out the importance of learning the terms, concepts, basic processes, and ordering options that would help the subjects in other *Advantage* courses and in their work-related functions. She briefly informed the subjects that the development team was investigating different methods of presenting this training, and that each learner had been assigned to one of three versions of it. She encouraged the learners to try their best throughout the program, and then gave a short demonstration of important key features within the CBT.

The learner initially entered demographics data that included highest education level attained, ethnicity, age-range, and gender that were stored by the computer for later use. After each learner submitted this demographic information, the course management system automatically opened the appropriate treatment version for that particular learner. Then, each learner independently worked through the CBT at his or her own pace, completing all modules in order. At the end of the last instructional module, the course management system directed the learner to the posttest and then to the learner attitude survey, both of which were completed on-line.

Criterion Measures

The posttest contained 40 multiple-choice items administered by the computer at the end of the training program. The posttest was identical for all subjects in all three program versions, and was based on the non-referenced version with no other-referencing or self-referencing. The computer program scored each item's response based on an answer key stored in the computer system. The posttest had been extensively field tested and revised in an earlier study (Fischer & Ku, 2000). The split-half Spearman-Bowman reliability coefficient for the posttest was .82.

En route scored quizzes were computer administered at the end of each of the five modules. In total, there were 40 multiple-choice en route scored-quiz items with each having four possible answer choices. Each correct answer was worth one point for a maximum possible score of 40 points. The en route quizzes were constructed from blueprints or tables of specification so that they would have the same total number and types of items per objective as the posttest. The reliability from the Spearman-Bowman split-half coefficient for the 40 en route items was .78.

Learners completed an end-of-instruction computer-administered attitude survey that consisted of 20 items that were identical across all versions with scores measured on a five-point Likert-type scale with 5.0 being the most favorable and 1 being the least favorable. Of the 20 items, 15 assessed the learner's levels of agreement with statements about elements of the course, including overall satisfaction, expected use on the job, pacing and directions, as well as features such as review and practice feedback and 5 items related to the learner's attitudes towards reading in general and to reading from computer screens. The alpha reliability for these 20 items was .93. There were also three constructed-response items related to items liked the best, least, and general comments, as well as four hypothetical items related to learning and motivation using alternative versions of the program.

The amount of time was measured for the overall course, each module, and for each type of screen (instruction, review, practice, quiz, etc.). All time was captured in seconds and later converted to minutes for ease in reporting.

Qualitative data were collected from 12 subjects (15%) in interviews conducted by the experimenter immediately following completion of the training. Of the 12 interviewees, five each had completed the other-referencing version and self-referencing version, and two had completed the non-referenced version. The interview protocol contained questions regarding CBT use, the best way for participants to learn, and subject's satisfaction with his or her own score, and with the training versions. Interviewees were also shown examples from the other versions and asked preference questions about them. Posttreatment interviews were analyzed for overall patterns in learners' perceptions. Responses to interview questions with yes/no components were summarized for frequency of responses by each treatment version.

Data Analysis

This study involved use of a posttest only experimental design with random assignment of subjects to the three treatments. Analysis of variance (ANOVA) at the .05 level was used to analyze the data for achievement on the posttest, en route quiz scores, overall time spent in the CBT and the time by module. Attitude survey data were analyzed using a multivariate of analysis (MANOVA) at the .05 level followed by univariate analyses where appropriate.

Results

The results are reported in this section for posttest achievement, en route quiz achievement, participant attitudes, and time in program.

Achievement

The posttest achievement data in Table 1 show that the mean scores for the three program versions were 27.52 (69%) for the non-referenced version, 31.32 (78%) for the other-referenced version, and 29.00 (73%) for the self-referenced version. The overall mean score across all contexts on the 40-item posttest was 29.27 (73%).

A one-way ANOVA of the posttest scores yielded a significant effect treatment versions, $F(2, 75) = 3.77$, $p = .03$, $\eta^2 = .09$. Follow up univariate analyses revealed that the mean score of 31.32 for the other-referenced group was significantly higher than the mean of 27.52 for the non-referenced group. The mean score of 29.00 for the self-referenced group did not differ significantly from that of either the other referenced or non-referenced groups.

Table 1 *Mean Posttest Scores for the Treatment Conditions*

Context	M	SD
Non-Referenced	27.52	5.88
Other-Referenced	31.32	4.36
Self-Referenced	29.00	4.48
Total	29.27	5.11

Note. Total possible score equals 40 items correct.

En route quiz scores

Mean scores for the 40 en route quiz items were 28.08 (70%) for the non-referenced version, 30.76 (77%) for the other-referenced version, and 28.96 (72%) for the self-referenced version. A one-way ANOVA of the en route scores indicated that there were no significant differences between the three versions, $F(2, 75) = 2.58$, $p = .08$, $\eta^2 = .06$. The overall mean score across all versions on the 40 quiz items was 29.26 (73%), the same overall percentage that subjects had on the posttest.

Attitudes

The overall mean score across the 20 Participant Attitude Survey items that were worded identically across all contexts was 4.04, a favorable rating indicating positive agreement with statements about the instructional program. Participants rated 11 of the 20 items about the CBT program at 4.0 or higher on a 5.0 scale. Several of these statements were related to the overall worth of the CBT with participants indicating that "they learned a lot", that "the training was worth their time", and that they "would recommend the training to others". The highest-rated items were related to using the CBT and included self-pacing, time to complete the course, and ease of using the CBT. Two statements with identical means ($M = 4.47$) had the highest ratings for

Items 1-20 on the survey. They were: Item 7, "I feel that there was sufficient time for me to complete this course" and Item 8, "I understood the directions on how to use the computer-based training".

Subjects in the other-referenced group consistently responded more positively to 16 of the 20 items about the program and its features than did subjects in the self-referenced and non-referenced groups. Manova with follow-up univariate analyses for these items, indicated that participants in the other-referenced treatment group had significantly more favorable scores on Item 15 than those in the self-referenced treatment group, $F(2, 75) = 3.79$, $p = .027$, $\eta^2 = .09$. Item 15 stated, "I feel that the CBT Final Test was a good assessment of the topics that were presented" ($M = 4.36$ for other-referencing and $M = 3.82$ for self-referencing).

Time

The mean times by treatment were 136.26 minutes for the non-referenced version, 135.30 minutes for the other-referenced version, and 113.23 minutes for the self-referenced version. A one-way ANOVA yielded a significant overall difference in time between treatments, $F(2, 75) = 4.66$, $p = .01$, $\eta^2 = .11$. Follow up univariate analyses revealed that the SR group spent significantly less time in the program than both the OR and NR groups.

Discussion

The supposition that participants completing the training that used the story contexts would attain higher achievement over the non-story or participants held true in terms of mean scores for both story versions. Participants in both treatments that had a story (OR and SR) attained higher posttest mean scores than did participants in the non-story (NR) treatment. However, there was a statistically significant difference only for one story context, the OR treatment group, compared to the NR treatment group.

The OR version was presented from the third-person point of view by an unseen narrator who described what Bea was doing, such as which on-line documents she would complete to pay for airline tickets for Dr. Cleff to attend a Jazz Conference in New Orleans. The use of a character with the story presented from the third-person point of view may have afforded the participants an opportunity to become actively engaged in much the same way that they become absorbed in reading a story or watching action such as on a screen. This finding would be consistent with design guidelines that emphasize third-person narration of screen text (Alessi & Trollip, 1991; Fleming & Levie, 1993). A study with adult college students concluded that third-person narration was more effective in enhancing retention and positive attitudes than the use of several characters delivering their script in a first-person point of view (Cates, Bishop, & Hung, 2000).

In support of the idea that a third-person story may be superior for retention are the results, which seem to indicate that learners related to the main character Bea and her boss Dr. Cleff in the other-referencing context as if they were real people. In posttreatment interviews, several participants who completed the OR version expressed their views on the story and the characters. One interviewee explained, "This version (OR) is like martial arts telling a story or rhyme. I can associate better with a story. When the CBT mentioned org managers, I could picture Dr. Cleff in the Music department."

Weller (2000) in describing how narrative was used to provide context, structure and appeal in an entry-level computing course, states that, "the use of characters creates a narrative which can be motivating and interesting for the student." (p.8). This idea is similar to what seems to have occurred in the present study. That is, that a story seen through the actions of someone like the learner may increase the learner’s interest and motivation. More than one participant in the posttreatment interviews related how they visualized Bea and Dr. Cleff in their workplace setting. One interviewee explained her preference for the other-referencing version, saying, "I liked having Bea--it made it more reality-based. Bea motivated me better."

Further support for the idea that computer-based training programs like the one in the present study motivate learners was reported in a pilot study of a computer-based learning program that used scenarios with fictional characters to teach accounting principles to university students. When asked why they liked the computer-based learning (CBL) accounting program, these undergraduate students cited self-pacing, active learning, that CBL was more motivating than other learning methods, and that CBL held their interest (Tam et al., 1997).

Participants' overall opinions about training were very positive with 11 of 20 items about the CBT program rated at 4.0 or higher on a 5.0 scale. OR participants had the most positive attitudes, rating 16 of these 20 items at 4.0 or higher. Similarly, in an accounting CBL that used scenarios, undergraduate students rated 13 of 20 items at 4.0 or higher on a 5.0 scale (Tam et al., 1997). It appears that learners regard computer-based instruction that includes scenarios or stories positively.

Participants in the present study indicated that "they learned a lot", that "the training was worth their time", and that they "would recommend the training to others". These opinions were similarly reported in a study of an interactive multimedia business simulation, in which 89% of the participants found the training worthwhile, 85% would recommend the training, and 84% would use the program again (Klassen & Drummond, 2000).

Participants in the SR version completed the CBT in significantly less time than did participants in either the non-referenced or the other-referencing versions. This finding is similar to those described in the meta-analysis of Symons and Johnson (1997), who suggested that both semantic and OR contexts seem to require more processing time than SR contexts. SR appears to takes less time to encode and retrieve because participants attach the new content to a well-developed and well-organized existing structure. Other studies have also shown that self-referencing tasks do not take longer than semantic tasks (Keenan & Baillet, 1980; Kuiper & Rogers, 1979).

Future Research

Results of this study seem to indicate that it is important to design computer-based training for adults within a context that enhances learning and provides for high motivation. Using a story as a way of presenting training in CBT appears to be an effective means for both increasing achievement and sustaining motivation. While the posttest results were significant for the other-referencing treatment over the non-referenced treatment, Participants in the self-referencing version also attained higher achievement than the non-story participants. This seems to indicate that the story played an important role, but yet to be determined is whether the point of view in presenting the story could be equally as effective for self-referencing as other-referencing in other settings and populations.

Overall time for completing the program may also vary with context, as indicated by the results of this study, in which participants in the SR treatment group spent significantly less time to complete the program than did those in the NR and OR treatment groups. The finding for time has been supported in other studies that used self-referencing, but not with a CBT (Keenan & Baillet, 1980; Kuiper & Rogers, 1979; Symons & Johnson, 1997). The amount of time required to complete a CBT is an important facet for training managers in some workplace settings that may influence design and development considerations, especially in situations that require shorter training times.

In addition to the setting and usage described in this study, it may be beneficial to consider investigating and reporting on the use of stories in other educational and business settings, with different themes, and at different levels of learning for K-16 and beyond to include adults in training and other types of instruction.

Conclusion

This study differed from previous research on non-referenced, other-referenced, and self-referenced studies in four main ways: (a) subjects were adults in the workplace attending required training, (b) the amount of time participants could and did spend in the program and each of the modules was not limited, (c) the study was conducted in an authentic environment under normal conditions for this type of training, and (d) the length of this instructional program was much longer (averaging two hours) than previously reported OR and SR research studies.

As computer-based training and multimedia are increasingly used to deliver training, it is important that the training be developed in a manner than enhances learning and engenders high motivation. The use of a story as a framework for situating the training appears to be an effective means for both increasing achievement and sustaining positive attitudes. This investigation was the first to study the effects of presenting instructional content to adults in the workplace in a non-referenced, other-referenced, and self-referenced manner. Results of this study indicate that CBT that uses a story context, particularly one in which the story is presented through another person or character can be a powerful form of instruction. Stories as a means of learning are not just for children in the classroom; stories can be a very powerful method for teaching adults in the workplace.

References

Alessi, S., & Trollip, S. (1991). *Computer-based instruction: Methods and development* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Anand, P., & Ross, S. (1987). Using computer-assisted instruction to personalize arithmetic materials for elementary school children. *Journal of Educational Psychology, 79(1),* 72-78.

ASTD. (1996). *The 1996 American Society for Training and Development report on trends that affect corporate learning & performance* (2nd ed.). Alexandria, VA: Author.

Bellezza, F., & Hoyt, S. (1992). The self-referent effect and mental cueing. *Social Cognition, 10(1)*, 51-78.

Bower, G., & Gilligan, S. (1979). Remembering information related to one's self. *Journal of Research in Personality, 13*, 420-432.

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18(1),* 32-42.

Brown, P., Keenan, J., & Potts, G. (1986). The self-referent effect with imagery encoding. *Journal of Personality and Social Psychology*, *51(5),* 897-906.

Casey, C. (1996). Incorporating cognitive apprenticeship in multi-media. *Educational Technology Research and Development*, *44(1),* 71-84.

Cates, W., Bishop, M., & Hung, W. (2000). Characterization versus narration: Empirical investigations of drama's role in multimedia instructional software. Paper presented at the *Convention of the Association for*

Educational Communication and Technology. (22nd, Long Beach, CA, February 17, 2000).

Cordova, D. & Lepper, M. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88*(4), 715-730.

Davis-Dorsey, J., Ross, S.M., & Morrison, G. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology, 83*(1), 61-68.

Don, A. (1990). Narrative and the interface. In B. Laurel (Ed.), *The art of human-computer interface design* (pp. 383-391). Reading, MA: Addison–Wesley.

Dwyer, H. (1996). Effect of personalization on reading comprehension. (Doctoral dissertation, Arizona State University, 1996). *Dissertation Abstracts International, 59,* Z3810.

Fischer, K., & Ku, H. (2000). Formative evaluation round 2: Comparing results. Paper presented at the *Convention for the Association for Educational Communication and Technology*. (23rd, Denver, CO, October 25-28, 2000).

Fischer, K., Savenye, W., & Sullivan, H. (2002). Formative evaluation of computer-based training for a university financial system. *Performance Improvement Quarterly*, *15(1),* 11-24.

Fleming, M., & Levie, W. (Eds.) (1993). *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed.). Englewood Cliffs, NJ: Educational Technology.

Gathany, N., & Stehr-Green, J. (1994). Putting life into computer-based training: The creation of an epidemiologic case study. *Educational Technology*, *34(4),* 44-47.

Greco, J. (1996). Stories for executive development: An isotonic solution. *Journal of Organizational Change Management, 9(5),* 43-74.

Keenan, J., & Baillet, S. (1980). Memory for personally and socially significant events. In R. S. Nickerson (Ed.), *Attention and performance (Vol. 8, pp. 651-669)*. Hillsdale, NJ: Erlbaum.

Keenan, J., Golding, J., & Brown, P. (1992). Factors controlling the advantage of self-reference over other-reference. *Social Cognition, 10(1)*, 79-84.

Klassen, J., & Drummond, D. (2000). Human resources skills: Learning through an interactive multimedia business simulation. *International Journal of Educational Technology, 2(1)*, 1-22, Retrieved April 19, 2003, www.outreach.uiuc.edu/ijet.

Klein, S., & Kihlstrom, J. (1986). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General, 115*, 26-38.

Ku, H., & Sullivan, H. (2000). Personalization of mathematics word problems in Taiwan. *Educational Technology Research and Development, 48(3)*, 49-59.

Ku, H., & Sullivan, H. (2002). Student performance and attitudes using personalized mathematics instruction. *Educational Technology Research and Development, 50(1)*, 21-34.

Kuiper, N., & Rogers, T. (1979). Encoding of personal information: Self-other differences. *Journal of Personality and Social Psychology, 37(4),* 499-514.

Lopez, C., & Sullivan, H. (1992). Effect of personalization of instructional context on the achievement and attitudes of Hispanic students. *Educational Technology Research and Development, 40(4),* 5-13.

Lord, C. (1980). Schemas and images as memory aids: Two modes of processing social information. *Journal of Personality and Social Psychology, 38(2),* 257-269.

Lord, C. (1987). Imagining self and others: Reply to Brown, Keenan, and Potts. *Journal of Personality and Social Psychology, 53,* 445-450.

Maki, R., & McCaul, K. (1985). The effects of self-reference versus other reference on the recall of traits and nouns. *Bulletin of the Pyschonomic Society, 23(3),* 169-172.

McLellan, H. (1993). Evaluation in a situated leaning environment. *Educational Technology, 33(2),* 39-45.

Reeder, G., McCormick, C., Esselman, E. (1987). Self-referent processing and recall of prose. *Journal of Educational Psychology, 79(3),* 243-248.

Rogers, T., Kuiper, N., & Kirker, W. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35(9),* 677-688.

Smith, R., & Birney, R. (1997). *The interactive accounting lab,* (2nd ed.), Chicago, IL: Irwin McGraw-Hill.

Symons, C., & Johnson, B. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin, 121(3),* 371-394.

Tam, M., Wedd, S., & McKercher, M. (1997). Developing and evaluation of a computer-based learning pilot project for the teaching of holistic accounting concepts. *Australian Journal of Educational Technology, 13(1),* 54-67.

Weller, M. (2000). The use of narration to provide a cohesive structure for a web based computing course. *Journal of Interactive Media in Education 2000 (1)*, 1-18. Retrieved April 19, 2003, www-jime.open.ac.uk.

Is a Website or E-Learning Product Working Well?
How Many Users Should You Test?

Theodore Frick
Tyler Dodge
Xiaojing Liu
Bude Su
Indiana University

Abstract

How can one determine efficiently if an informational website or an e-learning product is working well? Relatively small numbers of the target audience are needed to improve a product during formative evaluation and usability testing as part of product development and revision cycles. However, during summative evaluation, how many subjects are needed to determine product effectiveness?

When investigating the number of subjects needed for usability tests, a Poisson probability model was found to be a reasonable fit to extant data (Nielsen & Landauer, 1993; Virzi, 1990, 1992). However, this model was chosen on the basis of the number of subjects needed to identify important usability problems with a product, not for determining its effectiveness. To determine if a Website or e-learning product is working well, we investigated the predictive validity of a discrete Bayesian decision model: the Sequential Probability Ratio Test (SPRT) -- originally developed by Wald (1947). Fifty-one people representing a campus community participated in a usability test of the university library online catalog search tool, and the results were analyzed post hoc with SPRT re-enactments to simulate sequential decision making after testing each subject. Across a range of parameters, the Bayesian SPRT reached the same conclusion as reflected by the entire sample with many fewer subjects, utilizing typically small Type I and II error rates. The study provides evidence of the usefulness of the SPRT decision model in situations where determination of effectiveness is the goal (product works well or not). The SPRT maximizes efficiency by testing only as many users as necessary to reach a confident conclusion.

Introduction

When investigating the number of subjects needed for usability tests, a Poisson probability model has been found to be a reasonable fit to extant data (Nielsen & Landauer, 1993; Virzi, 1990, 1992). In perhaps the most contemporary review of issues relating to usability testing, Turner, Nielsen, and Lewis (2002) identified two central concerns: the reliability of traditional testing procedures, and the validity of the traditional model of problem detection. More specifically, regarding the formula used for estimating problem detection, they questioned whether the probability of a problem being detected can be modeled fairly with a unitary probability value.

Moreover, this model was chosen on the basis of the number of subjects needed to identify important usability problems with a product, *not* for determining its effectiveness. The purpose of the current study is to offer an approach to usability testing utilizing the Sequential Probability Ratio Test (SPRT) to determine product effectiveness (Wald, 1947). Rather than testing with a predetermined sample size, SPRT analyzes the knowledge accumulating during testing to determine when to stop testing, significantly reducing the number of subjects required. Wald's sequential probability ratio test (SPRT) went beyond the work of Thomas Bayes, who was concerned about how decisions can be reached as evidence accumulates. Wald's SPRT gives us rules for when to stop collecting evidence and reach a conclusion. The SPRT also tells us the likelihood that we would be reaching a wrong conclusion. The SPRT was originally used for manufacturing quality control decisions, and was considered so important that it was classified as a defense secret by the U.S. government during World War II.

Usability testing traditionally serves one of two purposes, either formative or summative evaluation, and the contrasting goals of these two forms of evaluations are reflected in approaches to usability testing as either problem detection or determining effectiveness. Most of the literature concerns problem detection, and a central tenet is that, given enough users and evaluators, most if not all of a product's usability problems may be uncovered. Of course, when ungainly numbers would be needed, a balance must struck between investment in usability testing and returns on investment, that is, identified problems. Problem detection studies traditionally use the probabilistic Poisson model to determine the number of subjects needed.

$\textit{Uncovered Problems} = N\,(1 - (1 - \lambda)^n\,)$
N: total number of usability problems in the design
λ: proportion of usability problems discovered while testing a single user
n: number of subjects

Given an accurate probability estimate, this simple formula provides a fairly good prediction of the number of subjects needed to determine certain proportion of usability problems. Offering the first evidence supporting use of the model, Virzi (1992) found that observing four or five users would reveal 80% of a product's usability problems, but this estimate and a host of related issues have been actively debated over the last decade. The accumulation of literature relating to problem detection has raised doubts regarding the certainty of the "five users" rule, as well as bringing to light several previously unrecognized issues relating to usability testing, including the probability of error detection, the assumption of homogeneity among users, the inconsistency between evaluators, and the definition of the usability task.

The first central issue relates to the probability of detecting a problem during testing. An average value of between .30 and .40 was suggested by a number of studies (Nielsen & Landauer, 1993; Virzi, 1990, 1992) and, based on the cumulative binomial probability formula, led to the statement that testing only four or five users will uncover 80% of the usability problems. Indeed, the diminishing returns after testing five users, a rule-of-thumb popularized in Nielsen's (2000) online *Alertbox*, continues to gain acceptance. While the rule holds true for probabilities in that range, other studies suggest that the actual probability of finding usability problems may be considerably lower (Lewis, 1994), with the result that usability testing would require a significantly greater number of users. For the p value of .16 that Lewis found, fully twice as many users would be needed to find 80% of the problems. Further, though Virzi asserted that the more severe problems would generally be identified before those of lesser import, Lewis found no such correlation; indeed, findings by Spool and Schroeder (2001) likewise challenge Virzi's claim, indicating that testing with a small number of users could be problematic for products with potentially hazardous problems.

Not only challenging the accepted sample size, concern over the probability levels of error detection has brought other issues to the discussion of usability testing. To begin, Caulton (2001) concluded that the assumption of homogeneity among users—the equal likelihood of all users to encounter all problems—not only accounts for the discrepancy between Lewis and his predecessors but compromises usability findings based on the assumption. Virzi's (1992) binomial model, Caulton explains, assumes homogeneity among the subjects, who "must be equally likely to encounter *all* problems" (p. 2). By introducing two classes of usability problems (common and rare) into the model, Caulton duplicates Lewis' (1994) findings that rare problems are not likely to be detected with only five subjects. Moreover, Caulton shows that heterogeneous subgroups likewise create the need for increased numbers of users to detect the same number of usability problems. Further, Caulton's conclusion accounts for the assumption by Virzi (1992), uncorroborated by Lewis (1994), that the probability of detecting a problem is positively correlated to the severity of the problem: "it is possible that p and severity *were* correlated in Lewis' data, but that subgroups masked the correlation" (p. 6). In this way, the discrepancy between Virzi and Lewis may be explained, but only by introducing the complex issue of user group composition into usability testing.

The problems associated with the homogeneity assumption were also put forth by Woolrych and Cockton (2001), who, like Caulton (2001), challenged the validity of Nielsen and Landauer's (1993) formula supporting their claim that five users are enough to detect the majority of usability problems. First, through a discussion of statistical theory, the authors showed that the probability of errors being found may be much lower than is fixed in the formula. To demonstrate their claim, they cite Spool and Schroeder's (2001) study in which goal-oriented testing drove the probability much lower than Nielsen and Landauer's 31%. Then, citing their own study of heuristic evaluation, the authors show that the probability of error detection depends not only on the severity of the problem but on differences between users, the same issue explicated by Caulton.

Just as different users encounter different usability problems, so do different evaluators identify the problems inconsistently, a pattern referred to as the evaluator effect (Hertzum & Jacobsen, 2001; Jacobsen, Hertzum, & John, 1998). In these studies and others (Molich et al., 1998), it was found that even when employing similar evaluation methodologies to test the usability of identical products, evaluators differ in their assessment of which observations constitute usability problems. The subjective and inconsistent identification of problems, even when using such relatively strict usability evaluation methods as cognitive walkthroughs and think aloud procedures among experienced professionals, lead to inter-evaluator agreement as low as 5% to 65%. On the one hand, this suggests that testing with multiple evaluators will uncover more and more varied problems than with a single evaluator, and indeed, Jacobsen, Hertzum, & John (1998) note that "the effect of adding more evaluators to a usability test resembles the effect of adding more users" (p. 256). On the other hand, the disparity among evaluators problematizes the "apparent reality of usability improvement achieved through iterative application of usability evaluation methods" (Lewis, 2001, p. 346).

In an article cited above, Spool and Schroeder (2001) reveal a fourth issue central to the question of the number of users, namely the definition of the usability task. In contrast to Nielsen and Landauer's testing with clearly defined tasks, or what Hudson (2001) calls "task-directed" testing, Spool and Schroeder allowed users to define their own goals, or "goal-directed" testing. That is, the five-user rule relates to situations in which all users engage in the same tasks of the product under evaluation, but when testing entails authentic users engaged in authentic tasks, the probabilities of error detection may be no higher than .16; at such low levels, the number of users Spool and Schroeder found necessary may range from around six to over thirty. Task-directed testing cannot achieve the coverage that goal-directed testing does, and authentic website use certainly entails a number of personal decisions, but the author's methodology directed the users to conduct a purchase which arguably does not fairly characterize the majority of tasks, even on commercial sites.

While a major goal of usability test is to identify design problems and recommend changes for a certain product, it is equally important to verify whether a product is working well enough that there is no need to invest additional resources on re-design and further evaluation; moreover, in the case of product effectiveness, we concern ourselves only with the effective use of the product, not the insights coming from testing for further development. While a considerable number of research studies address the identification of usability problems, our research team was unable to identify any significant literature addressing the number of users needed to conclude if a product is working well enough to stop further testing. Perhaps the simplest and most intuitive method is a simple calculation of success rate, or the percentage of successes encountered during usability testing. As Nielsen (2001) explains, success rates "provide a general picture of how [a product] supports users" and represent "the bottom line of usability" (n.p.), but beyond an explanation of the usefulness of tallying partial successes, he does not discuss such implications as the statistical limitations of such a metric.

The purpose of the current study is to employ the Sequential Probability Ratio Test (SPRT) (Wald, 1947) to determine the number of subjects needed to conclude whether or not a website meets a given effectiveness criterion threshold. Under the framework of classical hypothesis testing, the number of subjects needed to test a product's effectiveness can be predicted with the specification of acceptable levels of Type I and Type II errors, along with the population variance. Significantly, a relatively large sample size is usually needed to conclude whether the findings are generalizable. In this study, we propose an alternative approach using Bayesian reasoning to determine number of subjects needed in usability testing. Wald's (1945) SPRT offers an elegant framework for making statistical decisions between different courses of action. Through SPRT, one can use prior probabilities to express the preference for one or the other action and determine how these beliefs can change based on the accumulation of knowledge from observed data (Wald, 1945). Wald (1945) claimed that using SPRT to make a sampling plan leads to an average saving of at least 48% in the necessary number of observations, compared with the classical hypothesis testing. Later Colton and McPherson (1976) similarly found that using SPRT can achieve potential economy by testing fewer samples than fixed-sample-size while still attaining the desired level of statistical significance.

SPRT is a methodology for deciding between two alternatives under sequential observations. Though not developed under the framework of Bayesian reasoning, SPRT can be regarded as an extension of Bayesian theorem with addition of stopping rules (Frick, 1989). The central tenet of Bayes' theorem is its likelihood principle: *posterior probability* is proportional to *the prior probability* multiplied by the *likelihood* of that alternative, which can be expressed as follows:

Posterior probability $\propto$ Prior probability × Likelihood

Likelihood is the conditional probability of the event when a particular alternative is true. Prior probability is people's prior knowledge about the probability distribution of the alternative before the observation; after the observation, people's beliefs in the alternative will change due to likelihood principle. Posterior probability is people's knowledge about the probability distribution of the alternative after observation (Schmitt, 1969). When the observations are made sequentially, the posterior probability of one observation becomes the prior probability of the next observation. When several alternatives involved, the Bayesian theorem can be expressed as follows:

If

i. Alternatives are mutually exclusive and exhaustive;
ii. Let $P_0(A_i)$ be the prior probability of A_i;
iii. X is the observation;
iv. $P(X \mid A_i)$ is the probability of the observation given that A_i is true.

Then the probability of A_i is

$$P(A_i \mid X) = \frac{P_0(A_i)\, P(X \mid A_i)}{\sum P_0(A_i)\, P(X \mid A_i)} \qquad (1)$$

Assuming we have to decide between two alternatives with error probabilities of α (Type I) and β (Type II), by using Bayesian reasoning to compute the posterior probability of the alternatives after each observation, at a certain point the results will be highly in favor of one alternative over the other. SPRT offers stopping rules that can be used to cease observation and reach a conclusion with α and β error tolerances. The rules are as follows:

Rule 1: Compute the ratio (PR) of the posterior probabilities of the alternatives. If PR is greater than or equal to $(1 - \beta) / \alpha$, then choose the first alternative;

Rule 2: If PR is less than or equal to $\beta / (1 - \alpha)$, then choose the second alternative;

Rule 3: If neither Rule 1 nor Rule 2 is true, then another observation is needed. After a new result obtained, then update the posterior probabilities and reapply the three rules.

Let us apply this rule in the context of the present study: we want to decide between the hypotheses that either the website is effective (I) or the website is not effective (II). In this case, we need to make a sequence of observations in the context of usability testing to determine which option to choose. After randomly selecting a subject from the population using the web site, we calculate the probability ratio, PR:

$$PR = \frac{P_{e0}\, P_e^{\,s}\,(1 - P_e)^f}{P_{n0}\, P_n^{\,s}\,(1 - P_n)^f} \qquad (2)$$

In Equation 2, P_{e0} and P_{n0} are the initial probabilities of effectiveness and non-effectiveness, respectively. P_e represents the probability of randomly selecting a subject that would complete the usability tasks if the website is effective, and P_n represents the probability of randomly selecting a subject that would fail in the tasks even if the website is effective; α is the error probability of concluding the website is effective when it is actually not effective, and β is the error probability of concluding that the website is not effective when it actually is.

In Equation 2, we can assume P_{e0} and P_{n0} to be equal so that they can cancel each other out in the equation. Assuming that *s* and *f* refer to the numbers of users who are successful or not at completing tasks, we have contextualized stopping rules as follows:

Rule 1: If $PR \geq (1 - \beta) / \alpha$, then stop the testing and conclude that the website is effective;
Rule 2: If $PR \leq \beta / 1 - \alpha$, then stop the testing and conclude that the website is not effective;
Rule 3: If $(1 - \beta) / \alpha < PR < (1 - \beta) / \alpha$, then randomly select another subject and test again, increment *s* or *f* accordingly, recalculate PR, and apply Rule 1 to Rule 3 again.

This kind of reasoning assumes a number of necessary prerequisites, some of which are likewise assumed in inferential statistics:

1. Observations must be independent. That is, the outcome of one observation should not influence the outcome of another.
2. Observations must be randomly sampled. Random sampling is necessary for generalizing results from sample to population.

In addition, we make the following assumptions:

3. Alternatives must be mutually exclusive and exhaustive. (Though the stopping rules of SPRT may be applied to more than two alternatives, in this study, we consider only two alternatives, namely effective or not effective.)
4. The conditional probability of each alternative must be specified.

Because of its potential, SPRT has been widely applied in industry to test the quality of manufactured products, and in education too, Bayesian procedures have been used in computerized adaptive testing (CAT) to make mastery and non-mastery decisions. Studies have shown that the SPRT can be successfully applied in CAT using item response theory (IRT) (Frick, 1989; Frick, 1992; Lewis & Sheenan, 1990; Reckase, 1994). Frick (1989) argued that though SPRT does not take into account variability in item difficulty, discrimination, and guessing factors, the decisions of mastery or non-mastery reached by SPRT in his study agreed very highly with those reached through administering the entire item pools to examinees. Frick concluded that because of its simplicity and practicality, SPRT offers a viable model to achieve reliable results in CAT, provided that the method is used conservatively (e.g., small error probabilities). Like determining mastery or non-mastery of a educational content area, the task of determining site effectiveness is a binary decision; moreover, the task of determining site effectiveness with the fewest subjects possible is similar to the task of determining mastery by sampling as few test items as possible, thus warranting our application of SPRT to usability testing. In this research, we seek to establish whether the SPRT has predicative validity in reaching reliable conclusions as to a website's effectiveness using as few subjects as possible.

Method

A total of 51 people 18 years or older participated in this study at a large mid-western university and its community. The subjects were recruited through a method of stratified convenience sampling. First, we identified five strata of users of the University library resources: undergraduate students, graduate students, faculty, staff, and non-university affiliated community members. In order to have a sufficient number of subjects for SPRT analysis, we determined that we needed about 50 subjects, and to obtain a sample corresponding to the population demography of the university, we determined to seek the following proportions: 33 undergraduate students, 11 graduate students, 3 faculty, 2 staff, and 2 community members. Subjects in all strata were obtained according to convenience (discussed below) in precisely these proportions, with the exception that one additional undergraduate subject was tested.

Among the research participants, 5 subjects reported that they used the university library's online catalog often, 18 subjects used it occasionally, 20 seldom used it , and 8 had never used it at all. In addition to self-reported usage of the online catalog, subjects were asked to report their confidence using other similar searches. Specifically, asked to respond to the statement "I am confident using search engines" in terms of a 5-point Likert scale, 15 subjects strongly agreed, 27 subjects agreed, 7 subjects were neutral or undecided, 1 subject disagreed, and 1 subject strongly disagreed with the statement of confidence.

This research involved testing the usability of the online catalog of the Indiana University libraries (IUCAT). Determining the success of IUCAT, while of interest to stakeholders in the site's usability, remained of secondary interest after our primary question regarding the applicability of SPRT in determining the number of users necessary to determine success. Accordingly, rather than engage in the more thorough but challenging style of testing involving goal-oriented tasks, we focused our testing on particular tasks addressed by the catalog.

In order to provide some empirical basis for our task selection, we consulted the documentation relating to the usability testing of another university's online catalog, namely the study conducted by the Institute of Museum and Library Services of the University of Texas at Austin (2001). In one phase of their study, the researchers conducted focus groups with volunteers recruited from the University libraries staff; nearly three-fourths of the volunteers, being librarians from the public services cluster, were asked to represent "those library users who are served by the Web site and with whom the professional staff has contact on a regular basis" (Institute of Museum and Library Services [IMLS], 2000b). These librarians were asked to "think of a task that you typically do on UT Library on Line" and to "briefly describe this task" (IMLS, 2000a), and we coded the list tasks to identify the most prevalent among them: finding details on a specific book, and finding materials on a specific topic, including searches of works by a given author.

From these categories we developed our tasks, which involved (1) identifying the most recent book in the library system written by a specific author, and (2) determining to which library or libraries a specific book belongs. These two tasks involve many of the same procedures as other tasks we did not test; they entail use of many of the same features of the site, and they require many the same skills on the part of the user. We believe, therefore, that these two tasks are representative of most if not all of the other tasks addressed by the online catalog, and so we operationally define the catalog's success in terms of typical users' successful completion of these two tasks.

Testing proceeded in the following manner. After identifying the campus buildings with the greatest number of computer laboratories available for student use, we visited the laboratories in their rank order, on different week days, and at various times of day. When the laboratories were crowded, we solicited students waiting in line; otherwise, we solicited them at their workstation, working systematically through the laboratory. No more than eight subjects were recruited from any single laboratory, and no more than ten on any single day. Faculty and staff were solicited in a similar manner: we identified the schools with the most students and visited the buildings on different days and at different times; we positioned ourselves at a haphazard location in the building and systematically solicited faculty and staff at their desks. The community member was chosen by convenience and tested at home. Two researchers from the team conducted each usability test, one facilitating the testing procedures, and the other recording observations regarding the subject's activities during the completion of the two designated tasks. In addition, the subjects completed a brief questionnaire of their computer experience and background information. Testing proceeded in this way until the target samples were satisfied. The majority of the computer workstations featured Windows operating systems, though a small number of Macintosh machines were also used in the testing; all of the testing employed the Microsoft *Internet Explorer* software browser.

The SPRT analysis proceeded as follows. First, using *Statistical Package for the Social Sciences* (SPSS) version 11.0, we analyzed the descriptive statistics relating to subject background information, time spent on each task, and successfulness of the usability tasks. Second, we used a random number table to randomize the record order of the usability test data. As mentioned above, we had collected data using from four to eight subjects from each computer cluster and had labeled the records chronologically; the purpose of randomizing the record order was to avoid possible bias relating to the data collection procedure. Third, the data records were individually coded as either success or failure based on how well the subject had performed the tasks: specifically, if a subject succeeded on both tasks, this counted as a success case, but if a subject failed on both tasks, failed on either one of the two tasks, or only partially succeeded on one or both of the tasks, we coded it as a failure case. Forth, we used the SPRT simulation coded by Frick (2001) to analyze how many subjects would be needed to conclude whether the online catalog is effective or not. Finally, changing various parameters of the SPRT, we compared the number of subjects needed to determine effectiveness reached by different criteria.

Results

We first defined the SPRT parameters as follows. If the online catalog website is effective, we would expect 90% or more of the users to succeed in the tasks set to them. If the success rate is 60% or less, we would conclude that the site is not effective. In other words, we are presented with two alternatives: (1) the website is effective, and (2) the web site is not effective. Stated as conditional probabilities, we have this:

Probability (success | website is effective) = .90 or higher {1}
Probability (success | website ineffective) = .60 or less {2}
(α=.05, β=.05)

The first randomly selected subject from the pool of 51 subjects did not pass the test (this person failed on the second task). Thus, to this point we have observed one failure and no successes. The results are summarized in Table 1.1. The posterior probability was inclined toward determining the site's effectiveness as a failure at a .80 probability level, while the posterior probability of success was only .20. Still, SPRT could not make a conclusion at this time.

Table 1.1 *SPRT Results after 1 Subject*

Alternative	Probability			
	Prior	Conditional	Joint	Posterior
Success	.5000 ×	.1000 =	.0500 / sum =	.2000
Failure	.5000 ×	.4000 =	.2000 / sum =	.8000
			sum = 0.2500	

The second randomly selected subject succeeded on both tasks, so altogether we have observed one success and one failure. The SPRT results are presented in Table 1.2.

Table 1.2 *SPRT Results after 2 Subjects*

Alternative	Probability			
	Prior	Conditional	Joint	Posterior
Success	.2000 ×	.9000 =	.1800 / sum =	.2727
Failure	.8000 ×	.6000 =	.4800 / sum =	.7273
			sum =.6600	

After this turn, the posterior probability for failure dropped from .80 to approximately .73, and the posterior probability for success increased from 0.20 to approximately 0.27. Similar steps were repeated until the 12th subject was tested. Through this round of analysis, still only one subject—the first—had failed the tasks; the remainder were successes. The SPRT results at this point appear Table 1.3.

Table 1.3 *SPRT Results after 12 Subjects*

Alternative	Probability			
	Prior	Conditional	Joint	Posterior
Success	.9351 ×	.9000 =	.8416 / sum =	.9558
Failure	.0648 ×	.6000 =	.0389 / sum =	.0442
			sum = 0.8805	

With this turn, the posterior probability for success had risen sufficiently to make a determination:

$$.956 / .0442 = 21.629 \geq (1 - \beta) / \alpha = 19$$

Accordingly, we aborted testing and concluded that the website is effective.

In summary, the SPRT analysis of subjects from our sample pool in a random order allowed for success to be determined with 12 subjects. This result, of course, reflects the input data, which included one failure case among the successes; significantly, the failure could have appeared in any position among the first eight iterations of the SRPT with the same result. Given the same parameters and without a failure case among the first 8 entries, SPRT would have reached the conclusion of overall success with only 8 subjects, and conversely, if all of the initial entries were failure cases, SPRT would have determined overall failure with only 3 subjects.

In order to test the predictive validity of SPRT, we analyzed the same random data set under different conditions. Table 2.1 lists the results of running SPRT while keeping Alpha and Beta error constant (α=.05, β=.05) but changing the success and failure rates.

Table 2.1 *SPRT Test Results with α=.05, β=.05*

Rate		Observation			
Success	Failure	Success	Failure	Total Users	Conclusion
90%	50%	8	1	9	Success
90%	60%	11	1	12	Success
90%	70%	17	1	18	Success
90%	80%	46	5	51	No conclusion

The results reveal that as the success and failure settings became closer together, more subjects were needed to reach the conclusion regarding website effectiveness (for example, if the failure rate increased from 60% to 70%, SPRT required an additional 6 successful users to conclude the overall effectiveness of the website); moreover, as the difference between the success rate and failure rate approaches zero, the number of users required to reach a conclusion becomes exponentially large.

On the other hand, we could keep the success and failure rates constant and reduce the Alpha and Beta error levels. Table 2.2 displays the results of running SPRT while maintaining Alpha and Beta error levels (α=.05, β=.05).

Table 2.2 *SPRT Test Results with Success Level=90%, Failure Level=60%*

Level		Observation			
α	β	Success	Failure	Total Users	Conclusion
0.05	0.05	11	1	12	Success
0.03	0.03	12	1	13	Success
0.01	0.01	15	1	16	Success

These results reveal that as the Alpha and Beta error levels were reduced, more subjects were needed to reach a conclusion regarding website effectiveness (e.g., to reduce the Alpha and Beta error levels from .5 to .1, SPRT required an additional 4 successful users to reach a conclusion). These results demonstrate that SPRT can provide a relatively predictable method of estimating the number of users needed to determine website effectiveness.

Success and failure levels like those above may be common in educational settings, such as computer adaptive testing, but in commercial and industrial contexts—especially ones with high stakes, as in medical and military production—both success and failure levels are likely to be much higher: effective manufacture of pharmaceutical products, for example, may be as high as 99%, while manufacture may be deemed unsatisfactory even at levels as high as 95%. SPRT can be applied in these contexts as well, though, as seen above, a significantly larger sample size will be required to reach a conclusion. As displayed in Table 2.3, running our sample (n=51) through SPRT with levels like mentioned above results in a determination of non-success. As the success and failure settings get closer, more samples were needed to reach a conclusion.

Table 2.3 *SPRT Test Results with α=.05, β=.05*

Rate		Observation			
Success	Failure	Success	Failure	Total Users	Conclusion
98%	90%	31	4	35	Non-success
99%	90%	24	3	27	Non-success
99%	98%	41	5	46	Non-success

Discussion

In this study of website effectiveness, the usability data were analyzed by SPRT to determine whether the site was successful or not. Across a range of parameters including error tolerance and thresholds of success and failure, SPRT reached the same conclusion as reflected by the entire sample set, but SPRT required fewer samples to do so. For example, when we continued applying SPRT to our entire set of 51 samples, we reached the same conclusion of success as reached with only 12 subjects. This shows that SPRT bears predictability and reliability in its determinations.

The study provides good evidence of the usefulness of SPRT in usability testing: in summative evaluations or situations where determination of effectiveness rather than error detection is the goal, SPRT provides a method of data analysis with considerable flexibility. Even when performing a single rather than sequential or iterative calculation of success, SPRT affords a simple and sound alternative to raw percentages or statistical procedures such as beta distributions, which are likely to require more users for the same error rates. Moreover, when used sequentially to analyze usability data, SPRT can provide determinations at a substantial reduction of number of users.

At first glance, the requirement for analysis parameters such as success and failure rates may seem to be a limitation of SPRT, but in fact, the beta distribution likewise calculates results according to a cutoff rate, which is, in effect, the average of the success and failure rates; indeed, it may be considered advantageous that SPRT allows success and failure to be specified independently. Accordingly, these parameters should not be considered a fallibility left to the discretion of the analyst but, rather, an opportunity for the testing to reflect the needs of the stakeholders.

While not bearing upon the usefulness of SPRT in usability testing, a few points regarding the actual testing deserve to be mentioned. First, the percentage of failures encountered during the study needs qualification. In several cases, despite the subject's entry of the correct information using the correct submission procedures (e.g., conducting a "title search" of "all libraries"), the server produced incorrect results, that is, results inconsistent with the results produced under the same conditions at other times; despite the fact that the subject used the online catalog in the correct manner, we tallied this as evidence against the site's effectiveness. Further, in most of the testing situations, the subjects experienced inordinate server delays in receiving results; many subjects interpreted this as an error on their part and returned to the search page to review their input, or repeatedly clicked the submit buttons, or in other ways disrupted the original usage scenario. In every case, we let the encounter proceed to its conclusion—often to success, however slow, but in several cases converting what would have been a successful case to one of failure to accomplish the task. Not only did the server, through its errors and delays, contribute to the number of unsuccessful searches, but our own criteria for success may be regarded as unduly stringent. Specifically, only if a subject succeeded on both of the tasks did we regard the case as a success; if the subject was successful on one task but only partially successful on another, we counted the entire case as a failure—a definition of success perhaps not reflective of the website owner's own, but one that ultimately provided data suitable to SPRT analysis.

A second consideration is the inconsistency of the appearance of the search page indifferent situations. Specifically, the html coding of the search page specifies that, in the drop-down list from which the user selects which libraries to include in the search, the default or selected option is "all campus libraries," meaning all libraries on the local campus but excluding all libraries on other campuses. In common settings, this default setting is used to guide the search, unless the user selects otherwise, but in the computer laboratories available for student use, this default is overridden: the browser instead presents "all libraries," that is all libraries on all campuses, as the default. This variation resulted in inconsistencies among results. Since the participants were solicited by convenience, their investment in the testing was likely only casual, and indeed, while the tasks were commonplace, they were not intrinsic. As a consequence, though both of the tasks called for the subject to find a reference from any of the libraries within the university system, one of the tasks addressed an item located only in an off-campus library. On this task, then, users at computers other than the campus laboratory workstations would have had to change the option relating to library selection to retrieve the same results as users in the laboratories, that is, to find the correct reference; otherwise, a different result would consistently be returned by the search engine. We considered this difference to be a limitation of the testing procedures (e.g., subjects recruited without compensation) rather than a limitation of the website (though the default option bears significant implications on the usability of the system), and accordingly, for users whose default setting covered only campus libraries, we accepted the alternate answer as correct. As with the errors discussed above, this limitation may bear upon accepting the findings as representative of the site's usability, but not upon the usefulness of SPRT procedures in usability testing more generally.

Finally, a related consideration is the limitation of generalizing from the usability tasks to the catalog search engine more broadly. While several of the features were not tested directly (e.g., searches for journal titles), we nonetheless consider them to be similar in presentation and functionally to the tasks covered by the testing. Accordingly, we may tentatively generalize the site's effectiveness on the tasks tested to reflect the site's effectiveness for the related tasks. Still, this step is problematized by the interaction between the tasks and the libraries searched, but again, this does not pertain to the SPRT analysis.

While the study offers data regarding the usability of a particular website search engine, and while the methods and usability results may inform future studies of website effectiveness, the chief contribution of this study is the demonstration of SPRT's application in usability testing. Further studies may likewise contribute to this body of knowledge through several avenues of inquiry: they may continue comparing SPRT to other statistical procedures to establish its benefits and limitations; explore the range of applications of SPRT to gauge its usefulness and flexibility; and establish methods of implementing SPRT during testing to determine when to stop testing. It is hoped that the present study demonstrates the promise of such pursuits.

References

Caulton, D. A. (2001). Relaxing the homogeneity assumption in usability testing. *Behaviour & Information Technology*, *20*(1), 1-7.

Colton, T., & McPherson, K. (1976). Two-stage plans compared with fixed-sampling-size and Walt SPRT plans. *Journal of the American Statistical Association, 71*.

Frick, T. W. (1989). Bayesian adaptation during computer-based tests and computer-guided practice exercises. *Journal of Educational Computing Research, 5*(1), 89-114.

Frick, T.W. (1992). Computerized adaptive mastery tests as expert systems. *Journal of Educational Computing Research, 8*(2), 187-213.

Frick, T. W. (2001). Making Decisions Using Bayesian Reasoning. Retrieved November 30, 2002, from http://education.indiana.edu/~frick/decide/start.html

Hertzum, M., & Jacobsen, N. E. (2001). The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*, *13*(4), 421-443.

Hudson, W. (2001). How many users does it take to change a web site? *SIGCHI Bulletin*, May/June 2001. Retrieved October 8, 2002, from http://www.syntagm.co.uk/design/articles/howmany.htm

Institute of Museum and Library Services (2000a). Evaluation of Effective Library Sites. *Characterization of the use and usability of a web-based digital library* (Usability Testing section). Retrieved October 6, 2002, from http://imls.lib.utexas.edu/usability/evalmit.html

Institute of Museum and Library Services (2000b). Report Covering Project Activities to 6/30/2000. *Characterization of the use and usability of a web-based digital library* (Reports section). Retrieved October 6, 2002, from http://imls.lib.utexas.edu/report/activities00-06-30.html

Jacobsen, N. E., Hertzum, M., & John, B. E. (1998). The evaluator effect in usability tests. In C.-M. Karat & A. Lund (Eds.), *Human Factors in Computing Systems CHI'98 Summary* (pp. 255-256). New York: ACM Press.

Lewis, C., & Sheenan, K. (1990). Using Bayesian decision theory to design a computerized mastery test. *Applied Psychological Measurement, 14*, 376-386.

Lewis, J. R. (1994). Sample sizes for usability studies: Additional considerations. *Human Factors*, *36*(2), 368-378.

Lewis, J. R. (2001). Introduction: Current issues in usability testing. *International Journal of Human-Computer Interaction*, *13*(4), 343-349.

Littlewood, B., & Wright, D. (1997). Some conservative stopping rules for the operational testing of safety-critical software. *IEEE Transactions on Software Engineering, 23*(11).

Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D. et al. (1998). Comparative evaluation of usability tests. In *Proceedings of the Usability Professionals Association 1998 (UPA98) Conference* (pp. 189-200). Washington D.C.: Usability Professionals Association. Retrieved October 8, 2002, from http://www.dialogdesign.dk/tekster/cue1/cue1paper.doc

Nielsen, J. (2000). Why you only need to test with 5 users. *Alertbox*, March 19, 2000. Retrieved October 8, 2002, from http://www.useit.com/alertbox/20000319.html

Nielsen, J. (2001). Success rate: The simplest usability metric. *Alertbox*, February 18, 2001. Retrieved November 13, 2002, from http://www.useit.com/alertbox/20010218.html

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of INTERCHI '93* (pp. 206-213). Amsterdam, The Netherlands: ACM Press.

Reckase, M. D. (1983). A procedure for decision making using tailored testing. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (pp. 237–255). New York: Academic Press.

Schmitt, S. A. (1969). *Measuring uncertainty: An elementary introduction to Bayesian statistics*. Reading, MA: Addison Welsley Publishing Company, Inc.

Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is no where near enough. In J. Jacko & A. Sears (Eds), *Conference on Human Factors in Computing Systems: CHI 2001 Extended Abstracts* (pp. 285-286). Seattle, WA: ACM Press.

Turner, C. W., Nielsen, J., & Lewis, J. R. (2002). Current issues in the determination of usability test sample size: How many users is enough? In *Usability Professionals' Association 2002 Conference Proceedings* (n.p.). Chicago: UPA.

Virzi, R. A. (1990). Streamlining the design process: Running fewer subjects. *Human Factors and Ergonomics Society 34th Annual Meeting* (pp. 291-294). Santa Monica, CA: Human Factors and Ergonomics Society.

Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, *34*(4), 457-468.

Wald, A. (1945). Sequential method for deciding between two courses of action. Journal of the American Statistical Association, 40(231), 277-306.

Wald, A. (1947). Sequential Analysis. New York: Wiley & Sons, Inc.

Woolrych, A., & Cockton, G. (2001). Why and when five test users aren't enough. In J. Vanderdonckt, A. Blandford, & A. Derycke (Eds.), Proceedings of IHM-HCI 2001 Conference: Vol. 2 (pp. 105-108). Toulouse, France: Cépadèus Éditions. Retrieved October 8, 2002, from http://www.cet.sunderland.ac.uk/~cs0gco/fiveusers.doc

A Web Simulation on Educational Change: Challenges and Solutions for Development

Theodore W. Frick
Kyong-Jee Kim
Rui Huang
Indiana University

Barbara M. Ludwig
QUALCOMM

Abstract

The Diffusion Simulation Game (DSG) was created to provide graduate students in the Instructional Systems Technology (IST) Department at Indiana University with an opportunity to play the role of a change agent charged with bringing about innovations in teaching methods at a fictional junior high school. Originally created by an IST faculty member in a paper-based format, the game consists of a game board and several information and feedback cards. To play the game, teams of three to four students work together to develop a strategy for persuading as many people as possible to adopt a particular innovation. This game has formed an integral part of a core, residential IST course since the 1970s. In 2001, the IST Department decided to offer this course for the first time online, and thus needed the paper-based version to be converted to a Web-based medium. We began by conducting a needs analysis of the stakeholders (course instructors and past students) for the Web-based DSG. Next we developed a rapid prototype (Web mock-up of what the game might look like in a browser) and conducted some usability tests in order improve the DSG design. After several design iterations, we then created a computer prototype using PHP (a Web programming language) and XML (Extensible Markup Language) with HTML to provide the interactivity of the game in which learners make decisions and experience their consequences. We also conducted two rounds of usability tests with target users to evaluate and improve the design of the working computer prototype. The final product was embedded into a distance education course after four months of development. We will discuss our initial evaluation results, and the overall impact of the online game on the distance students and the course in general. Finally, we will discuss ongoing maintenance and plans for the future of the DSG and its implications for developing a simulation to teach systemic change in education.

Introduction

The Diffusion Simulation Game (DSG) was originally designed for a graduate-level class in the Instructional Systems Technology (IST) Department at Indiana University. It was designed to provide graduate students with an opportunity to simulate the role of someone charged with bringing about changes in teaching methods at a fictional junior high school.

Originally created by an IST faculty member in the paper-based format, the game consists of a game board and several information and feedback cards. To play the game, teams of three to four students work together to develop a strategy for persuading as many people as possible to adopt a particular innovation (Molenda & Rice, 1979). This game has formed an integral part of a core, residential IST course since the 1970s.

In a typical setting, students played the game as a group with 3-4 students for duration of 1 hour. While playing the game the students work together to choose a strategy to persuade people to adopt an innovation and write down the strategies they used and the number of people who were successfully adopted the innovation. After completing the game, the instructor provides a debriefing session, in which students reflect on how the strategies they used affect the success and failure of the simulated diffusion of educational change.

In 2001, the IST Department decided to offer this course for the first time online. It was virtually impossible to use a paper-based DSG to teach distance students. Thus, paper-based version needed to be converted to a Web-based medium. The Web-based DSG came into use in 2002 after a four-month development process. The game is currently being used in the distance and residential versions of a core class of IST. In this paper we will describe the process of converting the paper-based DSG into a computer-based version. We'll also discuss the challenges that we encountered and the solutions we came up with to create this online simulation game.

DSG is now available on the Web at the following URL: http://www.indiana.edu/~istdemo/dsg/login.phtml. Currently an Indiana University network ID is required to log in to this Website. The development team is working to make it available to the general public as well. The general public is expected to be able to log in to this Website by obtaining a guest account from Indiana University's central authentication system in the near future.

Procedures

In spring of 2002, a development team was formed to develop online DSG. The team consisted of three graduate students in IST and the first author who is a faculty member of IST. The game was developed as a course project in an advanced production class. The development process followed steps outlined by Frick and Boling (2002) for effective Web instruction, an inquiry-based process.

Needs Assessment and Product Design Requirements

We conducted interviews with the stakeholders to identify the needs for the design and development of the Web-based DSG. This included interviewing the faculty member who originally developed the game and has been using it in the residential course, and the faculty/student team who was planning to offer the online version of the class that summer. From those interviews, we gathered information on the target audience and the context and requirements for learning. Stakeholders expected the game to be delivered via the World Wide Web (as opposed to CD-ROM). The instructors expected that a history of each student's game play be kept for use in debriefing. The DSG game typically takes more than two hours to play, so we needed to design it so that students could pause and resume the game from where they left off, even at a different computer. Also, since it was difficult for distance students to work as *a group* to play the game in real-time, the new online DSG was designed for a single player.

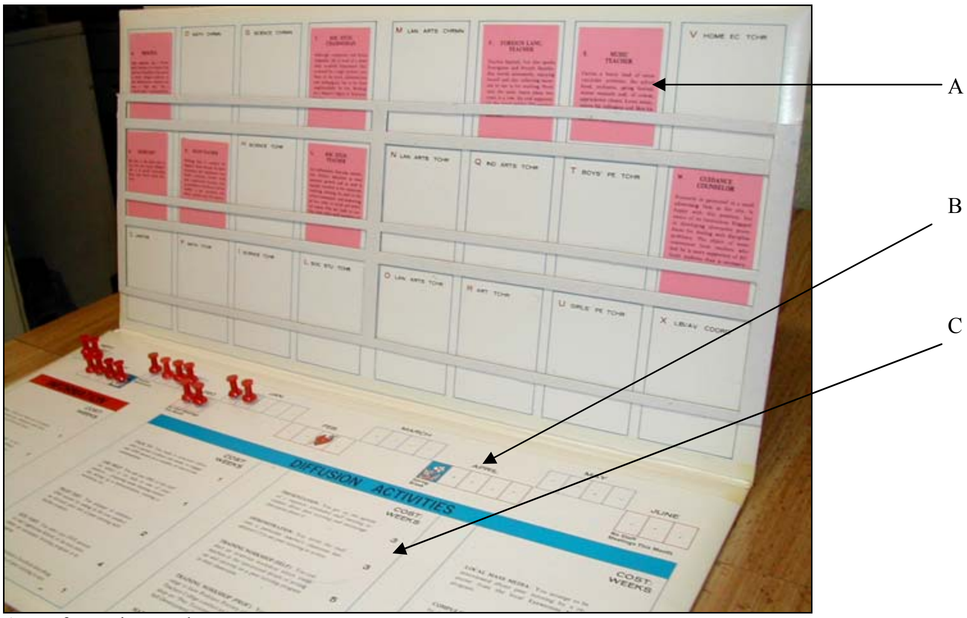
Design

Our aim was to transfer the paper-based game, which required a big table to place the game board, information cards, feedback cards, adoption form, instructions, and list of strategies onto a computer display whose size would be a minimum of 800 x 600 pixels. The main challenge of the interface design was to put the four major components of the paper-based version of the game (the game board, adoption form, information cards, and feedback cards) onto one dynamically changing Web page as the game progressed. A comparison between the structures of paper-based game and the Web-based one is shown in Table 1. Also, a picture of the paper-based game board and a screenshot of the online DSG page are shown in Figures 1 and 2.

Table 1. *Comparison between paper-based and the Web-based DSG*

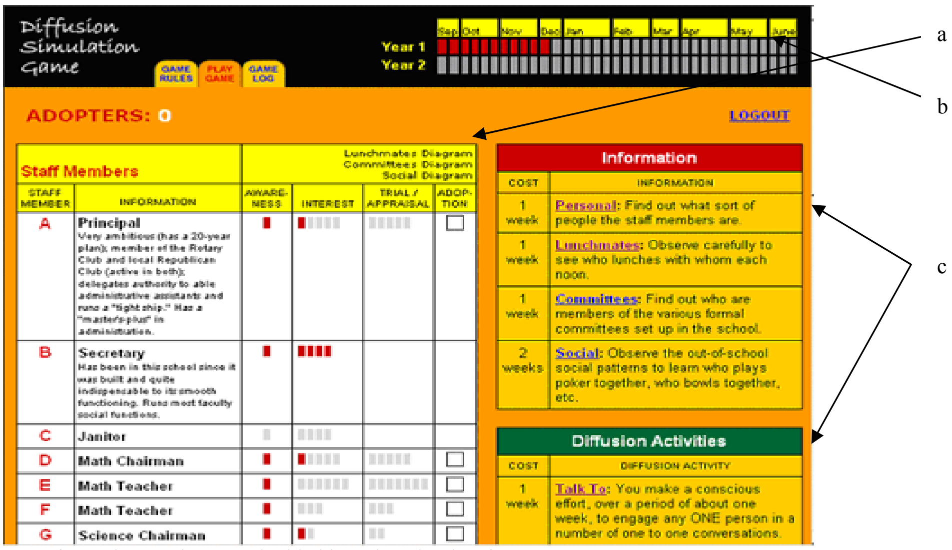
Components of paper-based game	Web-based game solution
1. Instructions for the game.	1. "Game Rules" page has detailed information about the game.
2. Game board – present available information and diffusion activities; store information cards; record passing time.	2. "Play Game" page – in which a calendar, characters list, information about strategies, and information gathered by the player are presented.
3. List of strategies – record strategies used and feedback on them.	3. "Game Log" page shows the history of the game, which includes information about strategies used and corresponding feedback.
4. Adoption form – record "influence points" for each character and adopters.	4. Adoption form is combined with the list of characters on the left side of the "Play Game" page.
5. Information cards – have detailed information about each character in the game.	5. Information about each character is stored in an XML file, which can be retrieved on the "Play Game" page after a player's selection.
6. Feedback cards – feedback on strategies users take from game facilitator.	6. Feedback is stored in XML files, which can be selected by the computer according to a player's decision and shown on "Play Game" page.

Figure 1. Paper-based DSG



A – Information cards
B – Calendar
C – List of information and diffusion activities

Figure 2. A screenshot of the "Play Game" page of the Web-based version of the DSG



a – Information cards are embedded into the adoption form
b – Calendar shown on the top of the page
c – List of information and diffusion activities

As the first step in this design process, we created a HTML-based mockup to design the interface of DSG. The concept of rapid prototyping (Boling & Frick, 1997; Frick & Boling, 2002) was used in this process as an iterative design approach. In this approach, the prototype is created on paper before it is developed on the computer so that the developer can test it out with the target users very early in the process. This iterative approach to design and development allows developers to evaluate and improve the design of the product early in the process. In this case, we created a mock-up HMTL Web page instead of the paper prototype due to fact that the game had been played on "paper" as a board game for many years in IST classes.

Then, we conducted several usability tests to evaluate and improve the design of the mock-up Website. The prototype was checked at various stages so that errors could be detected in the early stages. Usability testing is a process which involves observing the interaction of users representative of the target population with a product in order to identify and rectify usability problems (Rubin, 1994). Among existing usability testing protocols, we adopted the usability testing process by Frick and Boling (2002). We conducted usability tests of the DSG prototype with four graduate students in IST. Through this round of usability testing, some problems about the interface, diagrams, language are found and corresponding revisions were made of the HTML mock-up pages.

Development

A working computer prototype was produced in the development phase. For the paper-based game, a game facilitator is needed to give information game players require and feedback based on game rules to players' activities. Similarly, in a Web-based game, a server is needed to act as the facilitator and server-side programming is required. Another challenge was that unlike a paper-based game, which can provide several game boards to different player groups at the same time, there is only one game on one server. Therefore, data for different users had to be stored and differentiated.

To address those challenges of interactivity and data restoration, we decided to create recursive Web pages using PHP and XML. PHP (PHP Hypertext Preprocessor) is a server-side programming language that allowed us to create dynamically generated Web pages. Using PHP, it was possible to store the data of the game strategy, process it according to the game rules, and display the results on the Web page. Another reason that we used PHP instead of other server-side programming languages (e.g. ASP, JSP) is Apache - the server PHP works with most commonly - is free. Second, XML (Extensible Markup Language) is a document encoding or markup standard. We used XML to simplify and organize the structure of the feedback pages by storing and retrieving the information in extensible document elements. XML was also used to store user information in order to resume the game at any time from any computer.

Usability Evaluation

After developing the computer prototype, we demonstrated it to the stakeholders to solicit their feedback on the product. In addition, we conducted a second round of usability tests with target users. Four graduate students in IST at Indiana University participants in these usability tests. Here, we used the same protocol for usability testing as we did in the first round of usability tests. We discovered some minor programming errors and interface design problems from those tests and the prototype was revised to resolve those problems.

Implementation

After the four-month design and development process, the product was introduced to distance students in the summer, 2002. The distance students played the simulation individually and had a group discussion for debriefing using an asynchronous online communication tool. During the debriefing, they had access to their game history via log pages which used PHP and XML to track the different diffusion strategies and tactics they had used during game play.

The paper-based DSG, which had been used in a section of a residential core IST class, was also replaced by the online version 2003. The instructors took the advantage of this online game by not having to bring the bulky paper game sets to the class. Students also benefited from playing this game online, too. Distance students were able to play DSG at any place and at any time over the Web. Residential students seemed to enjoy playing this interactive online game, which was revealed by their reactions on playing the game from the evaluation data.

Evaluation

According to Kirkpatrick (1994), there are four levels of evaluation: reaction, learning, behavior, and results. For the DSG, we focused only on the first two, reaction and learning. The reaction level measures how participants react to an instructional program while the learning level examine the extent to which participants improve knowledge, skills, or attitudes as a result of participation. Level one evaluation has been conducted with users of the online DSG.

An online evaluation form was created to gather evaluation data from the students who played online DSG in a session of a residential IST core class which was offered in spring of 2003. DSG was also used in an online seminar on understanding systemic change in education during the summer of 2003 and evaluation questions were asked in an online discussion board. The results of evaluation will be discussed in more detail later in this paper.

Challenges and Solutions

There were several design and technical issues that the design team had to resolve to convert the paper-based game into an online version. These challenges and the solutions that the design team came up with are discussed below.

Screen Real Estate and Technological Constraints of Users

Our first challenge involved converting the paper-based game board to fit a computer display resolution of 800 x 600 pixels. Minimum technology requirements for the distance students at that time were a 56.6 Kbps modem internet connection, a sound card, and a computer display of 800 x 600 pixels. The paper version involved a lengthy instruction packet, a large game board, information and feedback cards, and forms for logging moves and recording points. We took several steps to reduce this to fit on a screen which included:

- Subsetting the game into three main Web pages: "Game Rules," "Play Game," and "Game Log," and including three tabs at the top of the interface so users can easily tab back and forth between these sections as needed.
- Aggressively editing the instructions and linking to non-critical information in pop-up windows.
- Linking to three diagrams used by players during game play in pop-up windows so that users can reference them easily without taking up screen real estate.
- Formatting the text in plain html (not Cascading Style Sheets). Because the game board includes a lot of text, formatting in plain html rather than CSS allows users to resize the text as needed via their browser settings.

Additionally, as the majority of our users were connecting to the Website via a dial-up modem, we had to ensure that the page sizes were small so that users could view and interact with the Website without excessive waiting. We knew that any frustration users may experience in accessing the game would have a negative impact on their learning experience. To ensure quick access, we avoided memory intensive graphics, and instead used PHP-generated HTML and static HTML as much as possible to create an attractive, easy-to-use, dynamically changing interface. We also streamlined and modularized our code so that pages would process as quickly as possible.

Comment: Do either of you know a citation or reference we could use for this?? How learner frustration diminishes learning?

Comment: I just try Google and find a paper at http://www.firstmonday.dk/issues/issue4_12/hara/ might be related to this, and the following paragraph is just about frustration and learning:
Sustained frustrations impede students' learning. Frustration interferes with pursuing goals (Reber, 1985) and thus it is one of the factors influence learning. Two aspects of learning are affected by frustrations: cognitive and affective (Jonassen and Grabowski, 1993). Research with college students shows that high levels of anxiety decrease the storage and processing capacity of working memory and impede making inferences (Darke, 1988a; Darke, 1988b). In addition, high frustration can demotivate students (Jonassen and Grabowski, 1993). Motivation is a strong factor that influences student learning (Alexander and Murphy, 1994; Covington, 1993; Stage, 1996). Especially, distance education requires that students be self-regulated (e.g. Abrahamson, 1998). In this kind of learning environment where students are away from traditional classrooms, frustration can be a major obstacle for distance learning.
And finally find one of the author is from IST ^o^

Comment: Rui, is this true? I know PHP isn't object oriented, but I imagine it is still possible to modularize and reuse code (?).

Barbara, I cannot say that because I don't know what "modularize" is exactly mean as a technical term. We indeed use some same codes in different pages, but I don't know if modularize is suitable for that.

Interactivity: Single vs. Multi-Player Game Play

The paper-based version is played in small groups whose members work together to choose strategies. Currently the online version does not allow for multiple people to play one game together at the same time. Early in the creation of the online prototype, we discussed adding this functionality, however, the game is already very complex even without the multi-player dimension. In the end, we determined that at that time, the benefit gained by making the online game multi-player did not merit the additional development time and potential for player confusion. The possibility of adding this functionality in future releases will continue to be evaluated as it may enhance the players' learning experience significantly.

Maintenance and Generalizability

The future maintenance and implementation of the game was something we had to plan for early on. Diverse sources outside of the department at Indiana University had already expressed interest in the game. There had been discussion of a corporate version of the game as well. We also had to consider the possibility that future DSG maintenance teams may not have extensive programming experience. We wanted to ensure the game allowed for easy maintenance and modifications to the interface, coding and feedback data.

There were several key development decisions made to achieve this end. We chose to program in PHP because it is a relatively easy-to-[illegible] language, and because it provided us the capability to code interactive, dynamically generated Web pages. Additionally we used XML to simplify and organize the structure of the feedback pages by storing and retrieving the information in extensible document elements. It allowed for an easy means for future modification of the feedback component of the game with a simple text editor.

Record-Keeping and Data Preservation

The DSG can take several hours to complete – especially when playing it for the first time. Players need the ability to be able to stop play and then resume it at a later time, possibly at a different computer. There is also an extensive debriefing exercise conducted with the entire class after everyone has played the game, which meant players needed access to their game history so they could review and refer to it during the exercise.

To make the game non-computer-specific meant we could not use cookies in our programming because they are stored on the client machine. We needed to store data on the server instead. We opted to use XML to save user information and game play in order to stop and restart the game as needed from any location. We also designed a "Game Log" page which implemented the XML technology in order to record all of the user's chosen strategies along with the corresponding feedback and results. They could then use this log to perform the debriefing exercise and compare their strategies and outcomes with other players.

Evaluation and Implications for a Simulation on Systemic Change in Education

The online version of DSG was first introduced to distance IST students at Indiana University in summer of 2002. In addition to the distance class, DSG has been used in two other classes as well: a core residential IST class and an online seminar on systemic change in education. Questionnaires were distributed to the participants of these two classes to assess users' reactions on DSG and the results are discussed below.

Evaluation: User Reactions

Approximately 30 graduate students in IST who were enrolled in a residential core class, Evaluation & Change in the Instructional Development Process, in spring 2003. They played online DSG in one class section as an activity to introduce them to diffusion and adopt of innovations. An online questionnaire was distributed to them to assess their reactions on playing online DSG.

The questionnaire consisted of 10 question items, which include the respondents' demographic information, the results of their game play. Eleven students returned the questionnaire and all the respondents were graduate students who were enroll in an IST core class at Indiana University. In terms of their performance, a majority of the respondents succeed in achieving the goal of the game, which was to obtain more than 10 adopters in a given period of time, in one or two attempts. Question items 4 and 5 included open ended question that asked the respondents whether they felt online DSG was realistic and what strategies they used to obtain adopters. Seven participants responded that they felt the game was realistic.

Deleted: 11

Deleted: 7

The question item 6-8 asked the respondents of their reactions of playing online DSG on a 5-point Lickert scale (1=strongly disagree, 5=strongly agree). The respondents agreed or strongly agreed that playing online DSG was a worthwhile learning experience (M=4.636, SD=0.504). The respondents also agreed or strongly agreed that they would recommend online DSG to others if they want to learn about the change process for adopting an innovation (M=4.545, SD=0.522). They also responded that online DSG motivated them to learn more about the change process for adopting an innovation (M=4.636, SD=0.50). When asked what they liked best about online DSG, the participants responded that the game was interesting because it was realistic, moderately challenging, and interactive.

Another evaluation was conducted of the participants of an online seminar on systemic change in education that was offered to in-service teachers in the summer of 2003. These participants played online DSG as an introductory activity to learn the topic of systemic change. After playing the game individually, they posted their reflections on their game play in an online discussion forum. 17 messages were posted on the discussion forum regarding their general reactions on the game play and whether they felt the game was realistic. In general, they exhibited positive responses on their experience playing this game. One participant stated:

> *First of all I would like to note how much FUN it was to play this game! My second immediate thought was what a great pedagogic tool for the classroom. You have to read, problem solve and think critically, many of the assets we want from our students in today's society. Too bad more simulation games like this aren't used in the classroom. I think students would find it fascinating, interesting and beneficial.*

Many participants also reported that the game was realistic. For instance, one participant stated:

> *I definitely think that it gives the participant a very realistic simulation of how innovations are adopted in a school system. First, I note the different roles and personalities within the organization and the complexity that lies therin.*

This simulation definitely highlights the importance of knowing who's who before you spend a lot of time on different diffusion activities. Position matters, but so does personality and social influence outside of one's formal role at work. I also thought that the two year timeframe was quite instructive. When you begin the game it feels like you have so much time that hitting your goal will be no problem but soon you realize how long it takes to win over adopters.

Change agents really need to have a reasonable long-term plan for implementing their innovations. Overall I found the simulation quite interesting and stimulating. It makes me think about my own school building and how I can introduce new ideas there.

Such positive responses from the users on the reality of DSG illustrates the possibilities to use Web-based simulation to teach the concept of systemic change in education, which will be discussed below.

Implications for Design of an Educational Systems Simulation on the Web: *SimEd*

DSG has demonstrated the potential for creating a simulation that provides interactive and prompt feedback to a player by taking advantage of current Web technologies. Based upon the experience of developing and implementing DSG, the first author expects to develop over the next several years a prototype simulation for designing educational systems that will run on the Web. He is calling this simulation, *SimEd*.

SimEd is envisioned to be something like *SimCity*, except that it will allow teachers, administrators, parents, students and school board members to design educational systems, not cities. As they design different variations of educational systems, they will see how a given design succeeds or fails over time. That is, they will be able to experience temporally the consequences of their design decisions, and they will be able to adjust their designs as their educational systems evolve. Thus, *SimEd* is expected to be a set of visioning tools. As an analogy, architects have created computer tools that allow them to create virtual buildings. These tools are very useful to envision what a particular design will be like, how it does or does not meet client needs, and to modify designs before committing to actual construction. Clients are able to virtually "walk" through such buildings before they are built, in contrast to looking at miniature models constructed out of balsa wood.

SimEd will help make more concrete some rather abstract and difficult-to-understand educational systems properties. But these properties are not those of physical buildings or cities. Rather, they will be based on the SIGGS Theory Model (Frick, 2002). The SIGGS Web site is merely descriptive of these educational systems properties. It does not allow people to experience them in a way such as *SimCity* provides an experience of designing a city and observing the long-term consequences. *SimEd* will use systems theory as a basis for the rules that drive the simulation through a period of compressed time – in particular, the educational systems theory proposed by Maccia and Maccia (1966). See http://education.indiana.edu/~frick/edutheo.html .

SimEd will utilize a combination of Web technologies. The computer interface will run on a user's Web browser using a plug-in for *Flash MX*. *Flash MX* will allow animation, drag-and-drop interaction, and playing of "movies" that are controlled by ActionScript, a Flash programming language. The Flash interface will interact with Web server programs written in PHP, a widely used programming language for Web applications. PHP scripts in turn will store system states and user data in XML files on the server. This strategy will allow SimEd to run on virtually any Web server and with any computer Web browser with the Flash plug-in installed (a free and fast download from Macromedia). The Diffusion Simulation presented in this paper is written in PHP and uses XML data storage. It does not use Flash MX, as MX was not available at the time. However, to get an idea of how these technologies can interact, see a drag-and-drop activity with feedback at: http://mentor.ucs.indiana.edu/~r641011/flashdemo/dragdrop.html . Indeed, the R641 class in spring 2003 demonstrated proof of concept that these technologies can work together successfully. See: http://www.indiana.edu/~tedfrick/r641/.

The *No Child Left Behind Act* of 2001 is likely to fuel the need for systemic change in education. *SimEd* is expected to help teachers, students, parents, administrators and school boards to: 1) understand the nature of systemic change in education; and 2) know what to look for in terms of consequences of new educational system designs.

References

Boling, E., & Frick, T. W. (1997). Holistic rapid prototyping for Web design: Early usability testing is essential. In B. Khan (Ed.), *Web-based instruction* (pp.319-328). Englewood Cliffs, NJ: Educational Technology Publications.

Frick, T. W. (2002). *Educational systems theory*. Retrieved September 29, 2003, from http://education.indiana.edu/~frick/siggs.html

Frick, T. W., & Boling, E. (2002). Effective Web instruction: Handbook for an inquiry-based process. Book in progress on the Web.

Kirkpatrick, D. (1994). *Evaluating Training Programs*. San Francisco, CA: Berrett-Koehler Publishers.

Maccia, E. S., & Maccia, G. S. (1966). *Development of educational theory derived from three theory*

models. Washington, DC: U.S. Office of Education, Project No. 5-0638.
Molenda, M., & Rice, J. A. (1979). The diffusion simulation game. *Simulation and Games, 10*(4), 459-467.

Rubin, J. (1994). Handbook of usability testing: How to plan, design, and conduct effective tests. New York: Wiley.

The Effects of Technology-Mediated Instructional Strategies on Motivation, Performance, and Self-Directed Learning

Donna M. Gabrielle
Florida State University

Abstract

The shift in education from an instructor-centered to a learner-centered focus requires learners to be self-directed and motivated. However, empirical data are lacking on how to positively effect self-directed learning. Further, the motivational needs of learners are often overlooked, and there is need for more literature examining both self-directed learning and motivation in technology-mediated learning environments. The purpose of this design experiment was to attempt to positively affect motivation, performance, and self-di rected learning of undergraduate students enrolled in a tuition-free, public military school. A second purpose was to use new technologies to deliver these instructional strategies as supplementary course conte nt. The subjects in this study were undergraduate students enrolled in a tuition-free, public military school in the Northeast United States. 784 students, representing approximately 20% of the population at the academy, were randomly divided into control and experimental groups for each of 16 instructors. Interventions were developed using Keller's ARCS model of motivation and delivered via PDA, web, and other technologies. The within-subjects research design used a mixed method approach involving both quantitative and qualitative data. Students completed four surveys to measure motivation and SDL. Course aggregate points were used to measure academic performance. The findings showed there were significant differences in motivation, academic performance, and proclivity to be self-directed learners of students who accessed the technology-mediated instructional strategies, suggesting that such motivational strategies should be incorporated into course instructional design.

Introduction

In the past century, the concept of distance learning has evolved from correspondence courses to instructional television to computer-based instruction to web-based learning. Today, the effort to put courses online is ubiquitous in education and training. Technology-mediated learning environments provide new opportunities for people to learn at their own convenience and pace. This shift in education from an instructor-centered to a learner-centered focus requires learners to be motivated and self-directed (Lee, 2000). However, empirical data are lacking on how to positively effect self-directed learning and satisfy the motivational needs of learners (J. Visser & Keller, 1990). Further, there is need for more literature examining motivation in technology-mediated learning environments.

Csikszentmihalyi, for example, spent over thirty years researching controlled experiments and case studies to investigate the importance of motivation and positive psychology (1990). His volumes of empirical evidence conclude that motivational issues are as important to learning as cognitive issues. Motivation has been found to be one of the most critical concerns in how and why people learn (Efklides, Kuhl, & Sorrentino, 2001; Keller, 1979). Keller (1999) notes that instructional designers are faced with even greater challenges in self-directed learning environments than with traditional learning, especially with regard to satisfying the motivational needs of learners.

Further, there has been some criticism about the quality of existing literature. Reeves (1995) analyzed decades of research studies in instructional technology (IT). He concluded that most studies in the field have problems including insufficient sample sizes, specification errors, lack of connection to theoretical foundations, meager treatment implementation, various measurement flaws, inaccurate statistical analyses, and futile discussions of results. The same author (Reeves, 2000, p. 4) said, "Given the poor quality of the inputs to research syntheses in the field of instructional technology, it is little wonder that the literature reviews and meta-analyses in IT yield disappointing results that provide practitioners with insufficient or confusing guidance." Similarly, Kulik and Kulik (1991) conducted an extensive literature review to examine the effectiveness of computer-based instruction. They argued that researchers conducting meta-analyses often reject more than 75 percent of published studies to include a small number that are worthy of additional analysis.

Thus, for researchers, the challenges include the need to pay particular attention to research design, development, and analyses. For instructional designers and educators, the challenges include the need to select the most appropriate instructional, technological, and motivational methods to improve learning. This study deals with some of these challenges.

Purpose/ Problem

Motivation is essential to learning and performance, particularly in technology-mediated environments where students must take an active role in their learning by being self-directed (Lee, 2000). However, empirical data are lacking on how to positively effect self-directed learning (SDL). Further, despite the importance of motivation to learning, J. Visser and Keller (1990) argue that the motivational needs of learners are often overlooked in educational research. To demonstrate this gap in the literature, L. Visser, Plomp, Arimault, and Kuiper (2002) examined the proceedings of the World Conferences of the International Council for Distance Education between 1988 and 1995. They found that less than one percent of the papers (only six of 801) focused on motivational issues. Consequently, more empirical research is necessary to examine motivation in technology-mediated learning environments.

Maslow (1970) defines motivation as a psychological process where a behavior is directed toward a goal based on an individual's needs. Keller (1999) argues that although motivation is idiosyncratic, learner motivation can also be affected by external aspects. These factors include systematic instructional design of tactics and strategies intended to improve motivation and performance, as well as encouragement and support by instructors, tutors, or peers. Thus, it would seem feasible that after conducting a motivational analysis of learners, appropriate strategies could be developed to improve motivation, performance, and SDL.

There is both experimental and correlational support for the instructional methods presented in this study. Song (1998) used Keller's Attention, Relevance, Confidence, Satisfaction (ARCS) model of motivation to develop computer-based instruction for middle school students. One control and two experimental groups received different levels of motivation in the instruction. Song found significantly higher levels of attention, relevance, motivation, and effectiveness in the group that received motivationally adaptive instruction than in the control group. J. Visser (1990) studied the impact of strategies designed using Keller's ARCS model that were delivered to adult learners. His embedded single-case exploratory study concluded that motivational messages could enhance learning by motivating students to undertake SDL tasks outside the classroom. J. Visser and Keller (1990) further studied the efficacy of motivational messages with adult learners in Mozambique, also with positive results. L. Visser (1998) took the concept of motivational messages a step further to encourage learners to persist in correspondence courses. She found that learners who received the motivational messages had reduced dropout rates and increased satisfaction. L. Visser, Plomp, and Kuiper later (1999) used similar strategies using the ARCS model with distance learners. In all studies, motivational messages were generally found to improve learner motivation, retention, satisfaction, and performance.

The purpose of this design experiment was to attempt to positively affect motivation, performance, and self-directed learning of undergraduate students enrolled in a tuition-free, public military school. A second purpose was to use new technologies to efficiently deliver these instructional strategies as supplementary course content.

Three questions were examined to achieve these purposes:

1. In applying the technology-mediated instructional strategies (TMIS) in the given instructional context, will a relationship exist between access to TMIS and academic performance as measured by course aggregate points (projects, homework, and examination grades throughout the semester)?
2. In applying the TMIS in the given instructional context, will a relationship exist between access to TMIS and proclivity to be self-directed as measured by the SDLRS instrument?
3. In applying the TMIS in the given instructional context, will a relationship exist between the access to TMIS and student motivation as measured by Keller's ARCS instruments (CIS and IMMS)?

Sample/ Population

The subjects in this study were undergraduate students enrolled in a tuition-free, public military school in the Northeast United States. 784 randomly selected students, representing approximately 20% of the population at the academy, agreed to participate in the study. The selected courses primarily consisted of freshmen and juniors in the graduating classes of 2003 and 2005. The courses were diverse in content (i.e. hard vs. soft sciences) to address the situational aspects of motivation.

Twelve courses were selected for the treatment and control, in a balanced design where each instructor had randomly assigned treatment sections and control sections. Within these courses, students in each section had identical syllabi and took identical examinations. Instructors were not informed of which sections were treatment and which were control.

Limitations of the Study

The most palpable limitation of the study was the homogeneity of the population. Prior studies in this publicly funded military academy reveal that the demographics of these cadets are not representative of traditional students (Hancock, 1991; Preczewski, 1997). For example, only 15% of cadets are female, all are unmarried with no children, and none are physically handicapped. The results are not likely generalizable beyond this context. However, this homogeneity also provides increased scientific control of the internal validity of the research.

Research Design

A mixed method approach involving both quantitative and qualitative components was employed in this study. Four surveys were used to measure motivation and self-directed learning: (1) the Course Interest Survey (CIS), developed by John Keller; (2) the Instructional Materials Motivation Survey (IMMS), developed by John Keller; (3) The Self-Directed Learning Readiness Scale (SDLRS), developed by Lucy Guglielmino, and; (4) The Self-Directed Learning (SDL) survey, developed for this study to track experimental group students' participation and perceptions of the strategies. The CIS, IMMS, and SDLRS were converted to web-based format and made available along with the SDL survey on the campus intranet. For all students in the study, academic performance (measured by course aggregate points including homework, projects, papers, and exams) was tracked throughout the semester, and extensive demographic data were also collected. See Figure 1 for the conceptual model.

Figure 1: Conceptual Model

Design of Technology-Mediated Instructional Strategies

Each TMIS consisted of three basic components (see Figure 2): (1) motivational messages at the beginning and end of each strategy; (2) supplementary instructional content, and; (3) the SDL survey to track participation and perceptions.

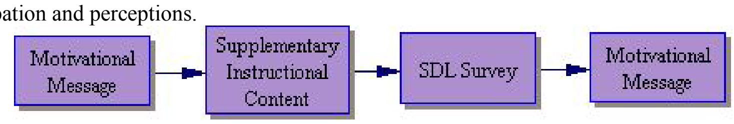


Figure 2: Systematic Design of Technology-Mediated Instructional Strategies

Keller's ARCS model of motivational design includes a ten-step procedure for instructional designers to develop motivational systems. The procedure was used for the design and development of TMIS in the present study and includes the following steps:

1. Obtain course information
2. Obtain audience information
3. Analyze audience
4. Analyze existing materials
5. List objectives and assessments
6. List potential tactics
7. Select and design tactics
8. Integrate with instruction
9. Select and develop materials
10. Evaluate and revise.

Most of the strategies were interactive instructional content provided for PDAs, on the Internet or academy intranet, or on CD-ROM. All technology-based content was provided in at least two formats (i.e. PDA and web) to ensure accessibility. See Table 1 for examples of ARCS components incorporated in the TMIS.

Instructional Content	Technology	ARCS Components
Instructional Resources	Web, streamed video, PDA	A, R, C, S
Self-Assessment	Intranet	C
Skill Briefs	PDA	A, R
Social Interaction	Threaded Discussion, IRC, Email	R, C, S
Mentor Support	Email	C

Table 1: *Examples of ARCS Components Incorporated in TMIS*

Figure 3 presents an example of a portion of a TMIS provided for PE210, Introduction to Wellness. ARCS components are seen throughout the design. The end of the content (not shown) supplies additional information which relates to cadets in their everyday lives and is intended to help students master course objectives. Figure 4 shows an example of the motivational messages portion of TMIS annotated with ARCS components.

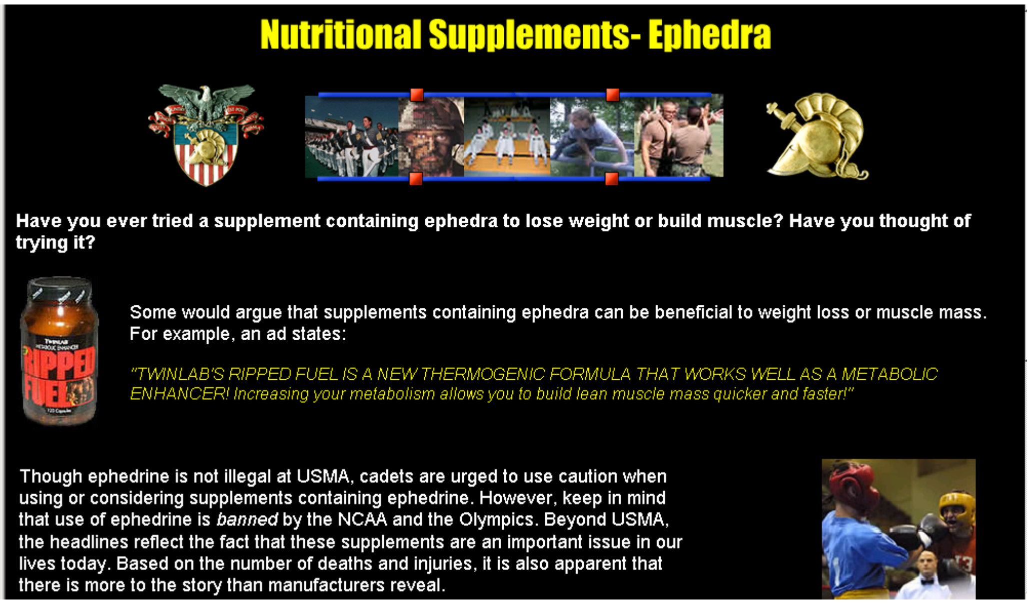


Figure 3: Example of Instructional Content Portion of TMIS

Hello again EN 102 Cadets,

Congratulations to all of you who participated in or supported the rigorous Sandhurst competition over the weekend! I also hope you did great on your last WPR. Thanks to all of you who have responded to the strategies so far (again, if you haven't, it's never too late). Most of you said you found the last strategy helpful, but as expected, some of you said you had too many other obligations to spend much time with the strategy. Based on your input, we are working to improve the strategies to try to make them as beneficial to learning the course material as possible.

You have been working on the challenging assignment of using your West Point experiences to produce a video that draws from *the Odyssey*. While not directly related to your video production, Strategy 6 should help you with your upcoming writing assignments, including HWE5. The strategy is provided in both web-based and PDA-based formats. Please give us your honest opinion on this supplementary material no later than COB 8 May. *Remember that even if you choose not to access the strategy, you should complete the brief questionnaire.* We would also strongly encourage you to provide your comments in the additional space provided on the web-based questionnaire.

1. Access **Strategy 6**: <> which is a study guide for the writing process. If you would like to add the web site to your PDA, click here to automatically add it to your existing AvantGo channels <>. Hint: when using your PDA on pages with lots of text, use the search function to quickly find what you need. Next, visit <> which is a study guide about writing correctly. To add the site to your PDA, click here to add it to your existing AvantGo account: <>.

2. Complete the brief online questionnaire: <> and please provide comments to give us the most beneficial feedback.

Thanks...and don't hesitate to contact me or CPT Corbett if you have any questions or if you need any assistance. Good luck on your HWE5!

Take care,

Donna Gabrielle

Attention- Text has good blend of colors and is easy to read.

Relevance - addressed to specific students in their own language.

Relevance - relates to their everyday lives.

Satisfaction- lets the students know that their feedback is being heard and that they are making a difference in the content they receive.

Relevance - letting students know that you are connecting what they have been working on to what they need to know for their next assignments.

Confidence - providing alternate forms of instructional content; both PDA and web.

Confidence- making the instructional content portion of the TMIS easily accessible and instructions on how to access the TMIS clear.

Relevance- using the name of their instructor and reiterating that you are working together to help them master the objectives.

Relevance - talking their talk. HWE5 means homework assignment number 5.

Figure 4: Example of Motivational Message Portion of TMIS

In order to respect student time constraints, the TMIS were delivered no more than six times per course. Student feedback allowed the instructional designer to use formative evaluation to improve the TMIS and document the changes throughout the process (see Figure 5). Customary of design experiments, the interventions are replicable beyond this study. The instructional content is specific to a course, though the TMIS can be re-used and modified as needed in subsequent semesters. The modifications are facilitated with the use of technology.

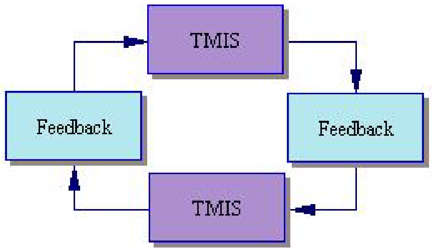


Figure 5: Formative Evaluation in TMIS Design Experiment

Procedures

The researcher met with academy department heads and course directors for input about the most appropriate courses to include. Course directors and department administrators selected instructors who reflected the diversity of all faculty including civilians and military officers. The researcher emailed all selected instructors to share information on their role in the study and request all course instructional materials so TMIS could be developed. A project website was developed to keep instructors informed. See Table 2 for a summary of the subsequent procedures and timeline.

	Control	Treatment
1. Visit classrooms in beginning of semester.	All	All
2. Email participants to thank them for participation.	All	All
3. Administer SDLRS pretest 2 weeks into study.	100 randomly selected	100 randomly selected
4. Administer CIS 4 weeks into semester.	All	All
5. Distribute TMIS a maximum of 6 times during semester.		All
6. Administer IMMS 2-4 times throughout semester.	All	All
7. Administer SDLRS posttest 2 weeks from end of semester to same students who took the pretest.	100 randomly selected	100 randomly selected
8. Collect performance data throughout semester.	All	All

Table 2: *Summary of Procedures and Timeline*

After communicating with instructors, the researcher visited all 48 classrooms to discuss the study and obtain signed informed consent forms. To prevent subject bias, the word "self-directed learning" was not used. A mock script was provided to instructors to ensure that students received the same information. Instructors encouraged participation by assuring students that the strategies would be brief, pre-approved by instructors, and designed to help them master course objectives. Students were assured that their participation was voluntary and confidential. All participating students signed informed consent forms before the study began.

An email message was sent to all participants thanking them for their participation. The SDLRS instrument was converted to web-based format and made available on the campus intranet. The SDLRS was called the Learning Preference Assessment to prevent subject bias. In the beginning of the semester, a random sample of 106 treatment group and 106 control group subjects received an email message directing them to complete the SDLRS. The rationale for limiting the number of students was the cost of the instrument. In the last two weeks of the semester, the same selected participants were asked to complete the SDLRS, providing a pretest/ posttest measure. It was assumed that four months between the tests was a sufficient interval to prevent test-retest issues. Some cadets were resigned from the academy before the end of the semester or did not complete the posttest, yielding 91 control and 104 treatment group students.

Treatment group students received TMIS via email. Each TMIS included motivational messages, a link to supplementary instructional content, and a link to the SDL survey. The SDL survey tracked participation and time on task, as well as open-ended questions for feedback about the TMIS. Control group participants did not complete the SDL survey since they did not receive the strategies.

Control and treatment group participants completed the online IMMS and CIS to quantify situational measures of motivation. The CIS was administered to all participants in the first two weeks of the study to measure their interest in each course. Email messages were sent to control group participants asking them to consider instructional materials related to specific course objectives to complete the IMMS. Email messages were sent to treatment group participants asking them to complete the IMMS to give their perceptions of the same instructional materials *and* associated TMIS.

Academic performance was measured for all participants by course aggregate points (examination, homework, projects, and other grades), except for EN302 students who were graded with pass/fail.

Data Analysis

Table 3 *Represents the hypotheses, instruments, and data analysis procedures used to test the hypotheses*.

Hypotheses	SDL	SDLRS	IMMS	CIS	Variables	Stat Test
H_{A1}: Treatment group students who used TMIS have significantly higher levels of academic performance (measured by course aggregate points including homework, projects, papers, and examinations) than control group students taught by traditional methods.	Usage				IND= Access to TMIS, instructor, course DEP= Academic Performance	General Linear Model, Resampling for WPPWE examination (P/F)
H_{A2}: The change in the mean pretest and posttest SDLRS score of treatment students with access to TMIS is significantly higher than the change in the mean pretest and posttest score of control group students taught by traditional methods.	Usage	Total Score (pretest/ posttest)			IND= Access to TMIS DEP= Δ SDLRS score	t-test, Wilcoxon Rank Sum test
H_{A3}: Treatment group students with access to TMIS have significantly higher mean vector scores for CIS and IMMS than control group students taught by traditional methods.	Usage		A,R,C,S, Total Score	A,R,C,S, Total Score	IND= Access to TMIS DEP= CIS score, IMMS score	Hotelling's *T*-squared test, MANOVA

SDL = Post-Strategy Self Directed Learning; SDLRS = Self-Directed Learning Readiness Scale; IMMS = Instructional Materials Motivation Survey; CIS = Course Interest Survey

Findings

Quantitative research findings indicated that there were significant differences in academic performance (p=.0045) between those students who accessed the technology-mediated instructional strategies and those who did not access the strategies. There were also significant differences in motivation as measured by the Course Interest Survey (p=.009) and the Instructional Materials Motivation Survey (p= less than .0001). Further, as hypothesized, the change in SDLRS scores was significantly greater for treatment group students than for control group students. There was a significant difference (p=.004) between the two groups, with control group cadets dropping by 3.3 points and treatment group cadets increasing SDLRS scores by 2.82 points.

Qualitative data showed that cadets felt the technology-mediated instructional strategies benefited their learning experience. For example, cadet comments included:

- "The tutorial was concise and quick to read... to the point and quick--just what cadets need."
- Seeing pictures to tie with the things we have heard about and were suppose to think about helped a lot. It made it more interesting."
- "The PDA is very helpful because you can always access your work and it has some interesting readings on it."
- "I am really enjoying the online discussions. In class, discussions are happening in real time, so a lot of very random thoughts are tossed out. However, on the web you have more time to think and contemplate what you are going to say before you say it. This ensures that the level of conversation

stays in the intelligent realm."

- "WOW!!! about sums it up. The link to the tutor pages was awesome. They had so much information that I had to go back a few different times this weekend just to see all that was there. This is good stuff, just like all that has been coming our way in order to help prepare us to be successful for the lessons and WPRs (exams)."
- "I really appreciated all of the extra study material all year, and I think that it should be made accessible and encouraged for all classes of basic chemistry."
- "The questions give the reader an idea of what to look for and also causes the reader to think more in depth on a subject or event he might have overlooked."
- "I enjoy the threaded discussions better than just writing out a journal entry. It helps to get different perspectives on the material from classmates as opposed to only hearing the instructor's side of the work."
- "It was fun!!! Made learning more interactive."

Qualitative data were very insightful as to how beneficial cadets perceived the strategies. Most were highly positive, though a handful of comments were negative, particularly with regard to frustration about using the technologies. Whenever a cadet made comments on the SDL questionnaire about problems with the technologies, the researcher contacted the cadet to help resolve the issue. Most problems were able to be resolved through email or phone dialogue, though a few cadets needed to bring their PDAs in to have the researcher fix them. Aside from that, many comments included wanting more interactivity and making the learning experience more fun.

Conclusions

This study revealed that systematically designed technology-mediated instructional strategies can be effective means of improving motivation, performance, and self-directed learning of students. Further, Keller's ARCS model is an effective method of developing such strategies and addressing the motivational needs of learners. Likewise, new technologies including PDAs can be efficient and effective means of delivering instructional content. Additional research is needed to show the effects of specific strategies on motivation and to test the instructional methods presented in this study with different populations.

References

Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. New York: Harper & Row.

Efklides, A., Kuhl, J., & Sorrentino, R. M. (2001). *Trends and prospects in motivation research*. Dordrecht; Boston; London: Kluwer Academic.

Gabrielle, D. M. (2002). *The Effects of Technology-Mediated Instructional Strategies on Motivation, Performance, and Self-Directed Learning* (Electronic): U.S. Military Academy Center for Teaching Excellence.

Gagné, R. M., Briggs, L. J., & Wager, W. W. (1988). *Principles of instructional design* (3rd ed.). New York: Holt Rinehart and Winston.

Guglielmino, L. M. (1978). *Development of the Self-Directed Learning Readiness Scale.* University of Georgia.

Hancock, D. R. (1991). *Effects of conceptual levels and direct and nondirect instruction patterns on achievement and motivation in course content (direct instruction).* Unpublished doctoral dissertation, Fordham University.

Keller, J. M. (1979). Motivation and instructional design: A theoretical perspective. *Journal of Instructional Development, 2*, 26-34.

Keller, J. M. (1999). Motivation in Cyber Learning Environments. *International Journal of Educational Technology, 1*(1), 7-30.

Kulik, C., & Kulik, J. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers and Human Behavior, 7*(1/2), 75-94.

Lee, C. Y. (2000). Student motivation in the online learning environment. *Journal of Educational Media & Library Sciences, 37*(4), 367-375.

Maslow, A. H. (1970). *Motivation and personality* (2d ed.). New York: Harper & Row.

Preczewski, S. C. (1997). *Measuring Self-Directedness for Continuing Learning: A Cross-Sectional Survey Approach Using the ODDI Continuing Learning Inventory (OCLI).* Unpublished doctoral dissertation, University of Missouri-Columbia, Columbia.

Reeves, T. (1995). *Questioning the questions of instructional technology research.* Paper presented at the Annual Conference of the Association for Educational Communications and Technology, Research and Theory Division, Anaheim, CA.

Reeves, T. (2000, April 27). *Enhancing the Worth of Instructional Technology Research through "Design Experiments" and Other Development Research Strategies.* Paper presented at the "International

Perspectives on Instructional Technology Research for the 21st Century," a Symposium sponsored by SIG/Instructional Technology at the Annual Meeting of the American Educational Research Association, New Orleans.

Song, S. H. (1998). *The effects of motivationally adaptive computer-assisted instruction developed through the ARCS model.* Unpublished doctoral dissertation.

Visser, J. (1990). *Enhancing learner motivation in an instructor-facilitated learning context.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

Visser, J., & Keller, J. M. (1990). The clinical use of motivational messages: An inquiry into the validity of the ARCS model of motivational design. *Instructional Science, 19*, 467-500.

Visser, L. (1998). *The Development of Motivational Communication in Distance Education Support*. Den Haag: CIP- Gegevens Koninklijke Bibliotheek.

Visser, L., Plomp, T., Arimault, R., & Kuiper, W. (2002). Motivating Students at a Distance: The Case of an International Audience. *Educational Technology Research & Development, 50*(2), 94-110.

Visser, L., Plomp, T., & Kuiper, W. (1999, February 10-14, 1999). *Development Research Applied to Improve Motivation in Distance Education.* Paper presented at the Association for Educational Communications and Technology, Houston, TX.

The Effects of Time-Compressed (accelerated) Instructional Video Presentations on Student Recognition, Recall and Preference

Joel D. Galbraith
Ying-Hsiu Liu
Bradley D. Ausman
Joshua A. Kirby
The Pennsylvania State University

Abstract

Anecdotal evidence and a body of research suggest increased learning as a result of moderately accelerated video instruction. This presentation discusses the results of a study exploring the effects and learning benefits of accelerated presentations in multimedia environments. Outside of the obvious time-savings benefits, implications are discussed for integrating accelerated video presentations into web-based and classroom instruction and its effects on recognition, recall and learner preference.

Introduction

A body of research exists (K Harrigan, 1995; K. Harrigan, 2000; King & Behnke, 1989; Olson, 1985; Orr, Friedman, & Williams, 1965; Short, 1977), suggesting effective learning can occur with time-compressed or accelerated video presentations. Leigh (1991)contends that “Parallel processing models acknowledge that a person can simultaneously engage in multiple tasks, such as seeing and hearing, depending on whether the tasks require the use of separate or common resources (, p. 72).” Outside of the, perhaps obvious, time-savings benefits, little is known about what type of learning objective or memory is affected by time-compression. This study sheds needed light on the issues, thereby helping instructors and instructional designers better understand time-compression and learning to aid them in creating more engaging and effective video-based instruction.

Time-compression and rate-modification in general, can be used in a variety of application areas including teaching, learning and human-computer interfaces. It has been used to skim media assets, speed up message presentation in voice mail systems and in aids for the blind and used to slow down media for learning languages, or for the hearing impaired (Aarons, 1992; Omoigui, He, Gupta, Grudin, & Sanocki, 1999). Over the last few years, schools and programs at many higher education and distance education institutions have incorporated streaming media in their instruction (Galbraith & Spencer, 2002). This use of streaming media has largely taken the form of online lecture archives and presentations as a primary means of instruction or in support of other modes of instruction. Large, distributed corporations and government agencies may well be leading the charge (Arlen, 2003; Galbraith, 2000), with online video presentations accounting for a significant portion of corporate training. Scores of inexpensive products and services exist in the market today that facilitate the simple and even fully automatic creation and distribution of rich streaming media and video presentations for education and training purposes.

Where instructional efficiencies in terms of time are desirable, the concept of time-compressed presentations has also long been an alluring aim in training and education. The functionality to accelerate media has only recently been made widely available as either player plug-ins (Enounce Inc.) or in Microsoft’s case, this variable speed playback functionality was just added to their most recent Windows Media 9.0 architecture that comes standard with Windows XP. With the increase in web-based video instruction and the growing simplicity with which time-compression can be applied to it, this study sought to better explain the effects of time-compression on memory and learning. What kind of learning does modest acceleration affect? (King & Behnke, 1989), and what kinds of learning tasks are most affected (Olson, 1985)? In particular, how might performance on cued recall and recognition tests vary under different time-compressed conditions?

Literature Review / Study Rationale

Moderately accelerated presentations have been shown to benefit learning (K. Harrigan, 2000; Short, 1977) as long as intelligibility can be maintained. In normal conversation, people speak and can comfortably hear words that are spoken at between 100 to 180 words per minute (Olson, 1985; Silverstone, 1974; Williams, 1998). When normal speech is increased to 210 words per minute, there is still no loss in comprehension (Omoigui et al., 1999). A 1965 study (Orr et al.) noticed that listeners could tolerate acceleration up to 2.0x normal speed (twice as fast), but they and others note that with ear training and practice, even higher speeds were possible (Aarons, 1992; Olson, 1985; Silverstone, 1974; Voor & Miller, 1965).

"This creates a speaking/listening discrepancy since a listener can comprehend spoken material up to four times (4x) faster than the speaker can send the message" writes Olson. "The result is a listener [or learner] who becomes bored or whose attention begins to wander" (1985, p. 3). Students themselves have reported anecdotally that accelerated presentations help them speed through boring or redundant material, it helps them stay more focused and attentive, learn more, and achieve higher grades (Galbraith & Spencer, 2002). Some research appears to bear out the aforementioned student comments under certain conditions (Gutenko, 1995; K Harrigan, 1995; K. Harrigan, 2000).

Not all the research is so positive, and time efficiencies are not always the prime concern at schools or even corporate environments. While media richness theory suggests than richer media can reduce ambiguity in conveying messages (El-Shinnawy & Markus, 1998), and can aid in converging on shared meaning, Information Processing (IP), and more specifically Dual Coding theory, question or temper the "more is better" approach. In her literature review of IP capacity, channels and modality, Hsia (1971) finds that an increase in the amount of information or stimuli presented, does not necessarily increase the rate of information encoding and storage. For example, much research suggests that intelligibility declines for users at some point of acceleration, directly compromising comprehension and memory (Beasley, Bratt, & Rintelmann, 1980; Heiman, 1986). In another example, King and Behnke's (1989) tests found that while short term and interpretive listening test scores (Kentucky Comprehensive Listening Test) did not decay until high levels of compression were reached, comprehensive listening scores and long term memory linearly declined with speed increases. This was attributed to the lack of processing time available for long term memory processing, and the rapid turnover of information in short term memory.

"Memory is commonly measured in two ways: recall and recognition. Recall is affected by almost all substructures of memory while recognition tasks typically entail a less rigorous involvement of memory structures" (Sundar, Narayan, Obregon, & Uppal, 1997, p. 2). For this study our intension is to follow the distinction posited by Sundar et. al. in our testing of memory.

The complexity of the stimulus materials and density of ideas presented are very relevant factors when it comes to time-compression and comprehension. IP theory holds generally that complex material would be adversely affected by high rates of time-compression due to overloading and subsequent over-writing of short term memory (Moore, Burton, & Myers, 1996). Along related lines, Limited Capacity models describe the allocation of resources (both voluntary and automatic) to given stimuli, Lang (2000) describes numerous conditions affecting orienting response and cognitive resource allocation relevant to video presentations, and how they are affected by the goals and needs of the individual—whether they are viewing for pleasure/relaxation or for learning. The highly attending learner-viewer is likely to run into resource-limited situations, but due to conscious increase in applied resources, the learner is still likely to process the message more fully (Lang, 2000).

Ear training, conditioning and acclimation to time-compression stimulus may also play a central role in how people respond to and learn from moderately accelerated presentations. Both empirical and extensive anecdotal evidence point to fairly rapid normalizing effects when listening to time-compressed speech. Beasley and Maki (1976), for example discuss how people's perception of what "normal speed" is and their preference for accelerated presentations, increase with prolonged exposure to time-compression. When accelerated presentations are returned to their recorded (normal) speed, people perceive the presentation to be artificially slowed down well below "normal" rates.

Those applying technologies to training and education have always sought for the perfect medium, one that delivers instruction so that learning is both effective and efficient (Fulford, 1993). Time efficiencies are at the root of much of the research in this field. In measuring time efficiencies, however, one cannot simply tally the length of the video presentations themselves. Students may pause and reflect more often and/or longer during accelerated presentations. According to Sticht (1969), listening to instructional material twice that has been accelerated by a factor of two, is more effective than listening to it once at normal speed. In such cases, obviously no time savings are realized. Similarly, Fulford (1993) reports that even with the ability to rewind and repeat—replay instruction, that accelerated presentations were more efficient.

Hypotheses

This study sought to better understand the nature or type of learning affected by time-compressed presentations. The main research questions (RQ) and Hypotheses (H) are as follows:
RQ1: For college students viewing instructional video presentations, controlling for speaking speed (wpm), presentation order and style, what is the relationship between time-compressed presentations and cued recall and recognition, where presentation speed and video segment are the independent variables?

H1: There will be no significant difference in performance on cued content recall tests between participants experiencing normal speed presentations and participants experiencing time-compressed presentations.

H2: There will be no significant difference in performance on content recognition tests between participants experiencing normal speed presentations and participants experiencing time-compressed presentations.

H3: There will be no significant difference in performance on cued context recall tests between participants experiencing normal speed presentations and participants experiencing time-compressed presentations.

RQ2: Do participants report a preference for time-compressed presentations over normal speed presentations?

Method

For the purposes of this study, Time-compression will be defined as the artificial modification of temporal media. "Normal speed" (1x), will refer to unmodified playback speed, and "Accelerated speed" refers to increasing playback speed by 1.5 times recorded or "normal" speed (1.5x).

The study was a 2x2 factorial design with 1 dependent variable (memory).

Video segment (IV)	Speed (IV)
Video segment V (Vollmer)	Normal speed
Video segment P (Peck)	Accelerated speed

.

Thirty-one (n=31) undergraduate Curriculum and Instruction (pre-service) education students enrolled in two "Internet for educators" classes were recruited for this study at a large eastern university. The two primary presentation video clips were selected specifically to cater to diverse educational interests and to strengthen the generalizability of the potential findings, while remaining relevant to the participants in their coursework. The two video presentations were comparable and controlled as to words per minute, but were qualitatively different in term of presentation style, topic, visualization, editing and variety of voices. Both video presentations were preceded with an introductory or "preliminary" video segment. The preliminary video's purpose (1 minute) was to provide accelerated group participants with time to acclimate to the time-compression condition.

Video segment "A" ran five minutes (5:00) under normal playback conditions and discussed the industrialization of the US public school system as the nation's economy evolved from an agrarian, resource-based economy to an industrial, skilled-labor based economy. Video segment "B" was narrated by multiple on and off camera presenters. The video was focused on outcome-based education and a related secondary education, home economics program in an eastern state. Both primary treatment clips averaged 163 words per minute (wpm) at normal speed, and all acceleration for the time-compression group was fixed at 1.5 times normal speed corresponding to an average of 244 wpm.

The independent variables employed in this study are speed of video segment presentation, and the video segment itself. The normal speed treatment group would view the two video segments in alternating orders and had a total running time of 11:00 minutes. The accelerated treatment groups viewed the same video segments at a time-compressed rate of 1.5x normal speed with a total running time of 6:36 minutes. The dependent variable employed in this study was memory. Specifically, queued content recall, content recognition and queued context recall. In addition, participant's preference for the different speed condition was sought. Participants were presented with both video speed conditions embedded side-by-side in a single webpage. During the course of this study the researchers specifically controlled for: participants ability to stop, pause, rewind or replay video clips and video segment presentation order. After completing informed consent forms, participants were randomly assigned to one of two instructional video presentation conditions (normal speed and accelerated). Participants were instructed to pay close attention to video presentations, as they would be quizzed following the treatment. All participants were pre-tested for prior knowledge of the treatment content and then proceeded to view their treatment materials.

Following the treatment video segments, the participants were administered a 36 item test. The test was composed of two series of 17 questions assessing "cued content recall" (7 questions), "content recognition" (6 questions), and "cued context recall" (6 questions). Upon completing the posttest, participants were prompted to view 2 short video segments; one playing at normal speed and one time-compressed and then indicate their speed preference.

Table 1

Normal	Instructions	Pre-test	Title (:06)	Preliminary video (1:00)	Title (:06)	Video V (5:00)	Title (:06)	Video P (5:00)	End (:06)	Memory Posttest	Total time 11:00
Accelerated (1.5 x normal speed)			Title (:06)	Preliminary video (:37)	Title (:06)	Video V (3:12)	Title (:06)	Video P (3:12)	End (:06)		Total time 6:36

When scoring the participants' responses to the assessment instruments, one point was awarded for each correct answer on recall and recognition questions related to the video segments' content and the video segments' context. All questions were fact-based questions derived from the video segments' content and context. All multiple-choice questions were objectively scored following a scoring key that was created by the researchers. A scoring rubric was devised to evaluate the short answer (cued recall) questions. Subjectivity in scoring the short answer questions was mitigated by using two independent graders. All discrepancies were reviewed by committee, negotiated, and ultimately scored and reported with full agreement.

A small sample of inferential statistics was used to analyze the data produced by the participants of our experiment. T-Tests, Pearson's correlation, and univariate analysis of variance (ANOVA) were used to compare the results between the treatment conditions. An alpha level of .05 was used for all statistical tests. For the pretest, T-tests were used to look for differences between the control (normal speed) group and the treatment (accelerated speed) group. The Pearson's correlation was used to test for a statistically significant relationship among the dependent variables. For the posttests, the analyses of the differences in memory performance utilized ANOVA and was used to test the treatment effects on the participants performance on the cued content recall, content recognition, and cued context recall test items.

Results

Descriptive Statistics

There were 30 students who participated in this study, with 16 participants in the control group (who viewed the normal speed presentation), and 14 participants for the treatment group (who viewed the time-compressed presentation). After reviewing the video segments, the students were asked to answer 12 cued content recall, 12 content recognition, and 10 cued context recall questions. For their memory level performance, please refer to Table 2.

Table 2 *Descriptive statistics for types of memory assessed.*

			Cued Content Recall		Content Recognition		Cued Context Recall	
Speed	Order	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Normal	0	8	4.63	2.92	9.50	2.07	3.38	1.19
	1	8	5.75	1.98	11.50	1.77	3.25	1.49
Time-compressed	0	6	5.33	2.42	9.33	2.66	3.00	2.10
	1	8	3.88	2.36	9.25	1.67	3.13	1.46

Note. For Order: 0 = Peck → Vollmer, 1 = Vollmer → Peck

Inferential Statistics

A pretest was administered prior to any treatments containing eight test items (five multiple choice and three short answer questions), and was designed to test the participants' prior knowledge of the content presented in the videos, including whether the participants had viewed the actual videos previously. The T-test was used to test whether treatment group and control group has equivalent performance in terms of prior knowledge and prior exposure to the videos. The results revealed that none of the participants were previously exposed to the video segments, and that there was no significant difference in the pretest performance between the control group (M = 1.81) and the treatment group (M = 1.86), t (28) = -.12, and p = .45 (two-tailed).

Manipulation checks

ANOVA was used to test whether presentation order played a statistically significant role across the four conditions. The data revealed that presentation order was not a potential variable and there was no significant difference in the mean score for the four groups in terms of cued content recall (F (28, 1) = .02 p = .90), content recognition (F (28, 1) = 1.47 p = .24), and cued context recall (F (28, 1) = .002, p = .96).

Pearson's correlation was also used to test the relationships among dependent variables across four conditions, and the data revealed that there is a statistically significant relationship between cued content recall and cued context recall (r = .442, sig.=.02). However, there is no statistical significant revealed between cued content recall and content recognition (r = .322, sig. = .083), and recognition and context recall (r = .026, sig. = .891).

An ANOVA using two independent variables, speed, video segment, and an interaction variable of speed and video segment, analyzed the impact on participants' various types of memory (cued content recall, content recognition, and cued context recall) across the four conditions (normal order 0, normal order 1, time-compressed order 0, and time-compressed order 1).

Cued content recall

H1: There will be no significant difference in memory performance in cued content recall between participants experiencing normal speed presentations and participants experiencing time-compressed presentations.

For cued content recall, ANOVA revealed that there is no significant difference in cued content recall across the speed conditions, F (1, 28) = .58, p = .14. Further analysis showed that the data did not reveal a significant difference in cued content recall in terms of video segments (F (1, 28) = 0.00, p = .50) and the interaction of speed and video segments (F (1, 28) = 0.00, p = .50).

The data does not reject H1; the participants experiencing time-compressed presentations did not exhibit significantly different performance in cued content recall tests than participants experiencing the normal speed presentations. See Table 3.

Table 3 *Cued Content Recall*

Source	*DF*	*DFDen*	*SS*	***F***	*P*
Speed	1	28	0.00	0.59	0.14
Video segments	1	28	0.00	0.00	0.50
Speed* Video segments	1	28	0.00	0.00	0.50

Recognition

H2: There will be no significant difference in memory performance in content recognition between participants experiencing normal speed presentations and participants experiencing time-compressed presentations.

For cued content recognition, again, we looked into whether speed, and video clips and interaction of these two variables had a significant impact on the overall cued content recognition performance of the participants.

ANOVA revealed that there is a significant difference in content recognition between these groups across the speed conditions (F (28, 1) = 2.51, p =.02). It was found that the participants of the normal speed condition scored significantly higher (M = 10.50, SD = 2.29) than those in the time-compressed speed condition (M = 9.29, SD = 1.91). In addition, the data showed that there was no significant difference in content recognition performance across the video segments and with the interaction of speed and video segments. Our data rejects H2 on the basis that one group scored significantly higher than the other group. See Table 3.

Table 3 *Content Recognition*

Source	*DF*	*DFDen*	*SS*	*F*	*P*
Speed	1	28	0.00	2.51	0.02†
Video segment	1	28	0.00	8.58	0.50
Speed* Video segment	1	28	0.00	0.10	0.50

† p is significant at the .05 level (2-tailed)

Cued context recall

H3: There will be no significant difference in memory performance in cued context recall between participants experiencing normal speed presentations and participants experiencing time-compressed presentations.

ANOVA revealed no significant difference between the speed conditions for the participants' cued context recall performance (F (28, 1) = 0.00, p = 0.50). Further analysis (ANOVA) revealed that for one of the video segments, there is significant difference in cued context recall performance (M_{Peck} = 3.20, M_{Vollmer} = 4.30, F (28, 1) = 6.4329, p < .05). This finding indicates that the "Vollmer" video segment was more conducive to cued context recall than the "Peck" video segment, which leads us to believe that further analysis on the characteristics of the Vollmer video should be conducted. The data does not allow us to reject H3. See Table 4.

Table 4 Cued Context Recall

Source	*DF*	*DFDen*	*SS*	*F*	*p*
Speed	1	28	0.00	0.00	0.50
Video segments	1	28	9.30	6.43	0.001†
Speed*video segments	1	28	0.43	0.30	0.22

† p is significant at the .05 level (2-tailed)

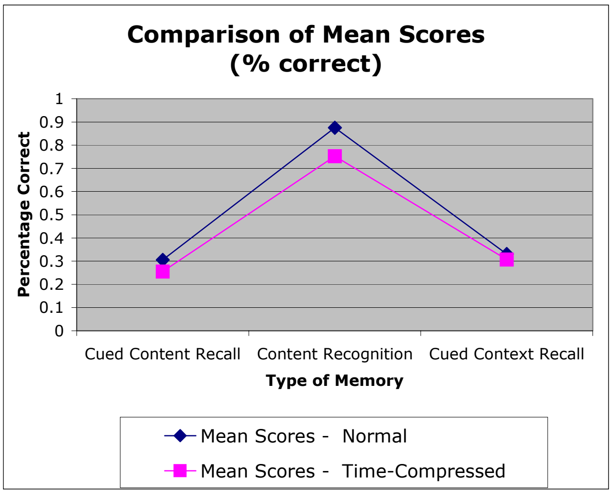
Presentation Speed Preference

In a surprisingly strong negative finding, a full 86.67% of students reported preferring normal speed presentation over the time-compressed presentation. The select few that reported a preference for accelerated presentations did not score any higher than those preferring normal speed presentations.

Summary of Key Findings

In summary, our data show a significant difference in the posttest scores for content recognition, which allows us to reject our H2 null hypothesis in favor of the better performance observed for the participants of the normal speed condition. Figure 1 shows the participants' performance on the posttests measuring the three types of memory tested across the speed conditions. The data does not support any significant findings for cued content recall or cued context recall that is in favor of either speed condition. There was a significant correlation of the participants' overall performance on the cued content recall and cued context recall. The participants' preference was overwhelmingly in favor of the normal speed condition.

Figure 1 Comparison of mean scores by memory type across speed conditions.



The difference between the participants' Content Recognition scores is significant.

Discussion

Figure 1 shows that the participants in the time-compressed condition did NOT score significantly lower than the normal condition for both of the cued recall measures. This is surprising because cued recall assessment items are generally considered more challenging due to the deeper levels of processing required. We might have expected the accelerated group to perform poorer as they had less time to process more information than did the normal speed group. A possible reason for the low overall recall scores may be due to test item difficulty. The test items looked for precise details found in the videos (e.g. exact words, phrases, contextual details), and such information may not have been effectively encoded by the participants and therefore difficult to retrieve. Points were not given for similar or near-correct responses, even if the participant indicated a basic understanding of the video presentations' messages. Anecdotally, this demonstrates that the participants were able to encode the general themes and concepts (or gist) of the videos, and such gist was not evaluated by the instrument. Assessing the effects of time-compression upon the generality versus the specificity of the participants' recall is one avenue for future research.

Our recognition data, on the other hand, showed significant differences in relation to speed – that is, the normal speed group outperformed the time-compressed group. Recognition assessment items are usually easier to answer since the participant receives multiple cues as to retrieving the correct answer from memory.

One possible reason why participants did not do as well in the accelerated condition may be due to a novelty effect and insufficient time to acclimate to acceleration. Tellis (1997, p. 76) argues that "when subjects first see novel stimuli, the novelty leads to uncertainty and tension." Due to the recent emergence and wide availability of this technology it could still be classified as quite novel. With regard to comprehension and acclimation time, Voor and Miller (1965) found that comprehension of compressed speech increased significantly over the first eight to ten minutes of listening with little increase after that. Our accelerated group experienced a total of only 7:01 minutes of time-compressed material—below the eight to ten minutes suggested acclimation time cited by Voor and Miller. This study's test instrument began testing material less than one minute in to the video segments, likely disadvantaging the acceleration group.

Numerous variables exist in the research that relate to acceleration and conditioning. Perhaps the most salient of these, is user control of playback speed or Variable Speed Playback (VSP). VSP ability allows user to dynamically accelerate video presentations to comfortable speeds for the given topic and users prior knowledge or familiarity with the subject matter. The superiority of user-controlled VSP over fixed rate acceleration is substantiated by multiple reports (Cohen, 2000; K. Harrigan, 2000; Omoigui et al., 1999; Short, 1977; Zemlin, Daniloff, & Shriner, 1968), but was not within the scope of this study.

Uses and Gratification theory may provide one possible explanation for the strong preferences for normal speed over accelerated. Additionally, traditional expectations of video being a more passive medium may be a variable. The expectation that video should require less cognitive processing than other forms of communication technologies may be directly challenged for students in the accelerated condition. Because their expectations/desires were not met, students may have reported their preference for normal speed. On the other hand, a more moderate acceleration of perhaps 1.3 times the normal presentation speed would likely have yielded more positive responses for acceleration. What is perhaps more relevant is that one's evaluation of what "normal speed" is, steadily increases with prolonged exposure to time-compression. The ultimate solution may lie in giving users individual and variable control over playback speed. This individualized ability has only very recently been made available in standard industry media players, and its effect has yet to be widely studied.

A primary motivation for time-compressed speech is for reducing the time needed for a user to listen to a message (Aarons, 1992). The findings of comparable performance in recall and recognition between the speed conditions seems generally to confirm previous research suggesting that learners can benefit from the time efficiency afforded by the accelerated speed of time-compressed presentations, with negligible impact on memory. The minor differences, however, that were observed in this study on the amount recalled between the two video segments suggests that the nature and/or the content of the video materials interacts with the viewers' interests and ability to remember the information presented. Future research should explore this possible interaction. Future research might also look at whether time efficiencies are in fact realized. Informal observations indicate that learners choose to view more content with acceleration, thereby defeating the potential time savings offered by time-compression. In addition, learners may be spend more time replaying specific segments of the videos in an effort to self-regulate their comprehension of the material being presented.

References

Aarons, B. (1992). *Techniques, Perception, and Applications of Time-Compressed Speech: Proceedings of 1992 Conference of the American Voice I/O Society.* Paper presented at the American Voice I/O Society.

Arlen, G. (2003). *For Tech's Sake: Video's Growing Role In Government Services*. Retrieved April 25, 2003, from http://washingtontechnology.com/news/1_1/for_techs_sakes/20605-1.html

Beasley, D. S., Bratt, G. W., & Rintelmann, W. F. (1980). Intelligibility of time-compressed sentential stimuli. *Journal of Speech and Hearing Research, 23*, 722-731.

Beasley, D. S., & Maki, J. E. (1976). Time- and Frequency-Altered Speech. In N. J. Lass (Ed.), *Contemporary Issues in Experimental Phonetics* (pp. 419-458). New York: Academic Press.

Cohen, G., Amir, A., Ponceleon, D, Blanchard, B., Petkovic, D., Srinivasan, S. (2000). *Using Audio Time Scale Modification for Video Browsing.* Paper presented at the 33rd Hawaii International Conference on System Sciences.

El-Shinnawy, M., & Markus, M. L. (1998). acceptance of communication media in organizations:richness or features. *IEEE Transactions on Professional Communication, 41*(4), 252-253.

Fulford, C. P. (1993). Can learning be more efficient? Using compressed speech audio tapes to enhance systematically designed text. *Educational Technology, 33*(2), 51-59.

Galbraith, J. (2000). *Using New Internet Technologies to Enhance the Effectiveness of Asynchronous Video-Based Instruction.*, Brigham Young University., Provo, UT.

Galbraith, J., & Spencer, S. (2002). *Variable Speed Playback and Video-Based Instruction: Is Faster Better?* Paper presented at the EdMedia, Denver, CO.

Gutenko, G. (1995). *Speed: "Run"- Time Compressed Video for Learning Improvement and Digital Time Compression Economy*. Retrieved March 24, 2003, from Eric Document, ED384 341.

Harrigan, K. (1995). The SPECIAL system: Self-paced education with compressed interactive audio learning. *Journal of Research on Computing in Education, 27*(3), 361-370.

Harrigan, K. (2000). The SPECIAL system: Searching time-compressed digital video lectures. *Journal of Research on Computing in Education, 33*(1), 77-86.

Heiman, G. W., Leo, R. J., Leighbody, G., & Bowler, K. (1986). Word Intelligibility Decrements and the Comprehension of Time-Compressed Speech. *Perception and Psychophysics, 40*(6), 407-411.

Hsia, H. J. (1971). The information processing capacity of modality and channel performance. *AV Communication Review, 19*(1), 51-75.

King, P. E., & Behnke, R. R. (1989). The Effect of Time-Compressed Speech on Comprehension, Interpretive and Short-Term Listening. *Human Communication Research, 15*(3), 428-443.

Lang, A. (2000). The Limited Capacity Model of mediated message processing. *Journal of Communications, 50(1), 46-70.*

Leigh, J. H. (1991). Information processing differences among braodcast media: review and sugestions for research. Journal of Adverstising, 20(2), 71-75.

Moore, D., Burton, J., & Myers, R. (1996). Multiple channel communication: The theoretical and research foundations of multimedia. In D. H. Jonassen (Ed.), Handbook of Research for Educational Communications and Technology. Mahwah, New Jersey: Lawrence Earlbaum.

Olson, J. S. (1985). A Study of the relative effectiveness of verbal and visual augmentation of rate-modified speech in the presentation of technical material. Paper presented at the Annual Conference of the Association or Educational Communications and Technology (AECT), Anaheim, Ca.

Omoigui, N., He, L., Gupta, A., Grudin, J., & Sanocki, E. (1999). Time-Compression: Systems Concerns, Usage, and Benefits. Paper presented at the CHI 99, 1999, Association for Computing Machinery, Pittsburgh, PA.

Orr, D. B., Friedman, H. L., & Williams, J. C. (1965). Trainability of listening comprehension of speeded discourse. Journal of Educational Psychology, 56, 148-156.

Short, S. (1977). A Comparison of Variable Time-Compressed Speech and Normal Rate Speech Based on Time Spent and Performance in a Course Taught with Self-Instructional Methods. British Journal of Educational Technology, 8(2), 146-157.

Silverstone, D. M. (1974). Compressed speech. audiovisual instruction, 19(1), 42-43.

Sticht, T. G. (1969). Comprehension of repeated time-compressed recordings. The Journal of Experimental Education, 37(4).

Sundar, S. S., Narayan, S., Obregon, R., & Uppal, C. (1997). Does web advertising work? Memory for print vs. online media. Paper presented at the 1997 Association for Education in Journalism and Mass Communication Conference: Communication Technology & Policy Division, Chicago.

Tellis, G. J. (1997). Efective frequency: One exposure or three factors? Journal of Advertising Research, 37(4), 75-80.

Voor, J. B., & Miller, J. M. (1965). The effect of practice upon the comprehension of time-compressed speech. Speech Monographs, 32, 452-455.

Williams, J. R. (1998). Guidelines for the use of multimedia in instruction. Paper presented at the Human Factors and Ergonomics Society 42nd Annual Meeting.

Zemlin, W. R., Daniloff, R. G., & Shriner, T. H. (1968). The difficulty of listening to time-compressed speech. Journal of Speech and Hearing Research, 11, 875-881.

Course Evaluation and Revision-Before, During and After: A Systematic, Collaborative and Effective Process to Improve Courses

Scott Hampton
Toya Davis
United States Military Academy

Abstract

Two instructors implemented a systematic process to evaluate and revise the actual design and delivery of Instructional Design courses at the undergraduate level. The process involved collaborative instructor assessments, observations and proposed revisions to develop a collaborative end of course evaluation with the students through focused discussion on tailored survey questions. The evaluative process, which could be applied to any course, survey questions, and pre-post intervention results are presented.

Introduction

"Please save us from ourselves. A discussion would be effective." Comments like this taken from a student end of course survey speaks directly to the need for change in the classroom. With a changing society and growing demands on time and resources, it is imperative that we continually search for ways to improve upon the old and revise when necessary. The classroom cannot be left behind in the process. Many instructors are therefore faced with the challenge of meeting the needs of a changing generation of students by revising instruction in an effort to improve upon the efficiency and effectiveness of both the instruction and the instructor.

The purpose of this study was to systematically evaluate and improve an introductory Instructional Design course using course evaluation principles and techniques expressed at AECT 2001 in Atlanta during a Presidential Session (Reiser & Hampton, 2001) on teaching. The presentation illustrated techniques for instructor self-assessments and reflections (Schön, 1987) of a course before the semester, before and after each class during the semester, and at the end of the semester. The instructors in this study formalized a systematic process to incorporate these 'continuous evaluation' techniques for four undergraduate sections over a two year period.

Prior to this evaluation, the course was designed with fourteen individual assessments in 39 lessons to include two exams, nine project performance submissions, and three presentations. This averaged to one graded event every 2.7 lessons. The systematic and progressive design of the course required that students correct mistakes in their project submissions so that they could succeed in the following steps. This model greatly limited the instructor's ability to provide timely feedback to the students for them to incorporate into subsequent graded requirements. In addition, the enrollment size of the course averaged 20 students per year. However, increased interest in the program drove the enrollment up by 50% to 30 meaning the demands on the instructor to provide timely feedback will significantly increase. Change was not only necessary but inevitable.

Method

Participants.

Four sections of undergraduate students (39 total) voluntarily participated in data collection procedures throughout 2 terms over a two year period. The instructors conducted a self-assessment of the course before the first term began by evaluating the course syllabus, learning outcomes and activities outlined by the previous instructor. Minor changes were made to the syllabus and learning activities to facilitate learner preparation and relevance of the course activities. The graded requirements were reduced to thirteen by removing one exam. After each class period (40) the instructors would reflect on the class (Ingram, 2001; Schön, 1987) in areas such as effectiveness of the learning activities, student comprehension/interest, and scheduling/assessment of activities. Additionally, one instructor observed the undergraduate course through a 1-way observation window to add another set of 'evaluation eyes' to the study.

Design

A pre-post quasi-experimental design examined the effects of course evaluation and revision before, during and after the course was taught. Changes were incorporated prior to the start of evaluation based on the feedback from the prior year (before). End of course feedback from the first year of evaluation was then compared to end of course feedback from the second year of evaluation (during and after).

Measures

During the last class meeting of each term, the instructors conducted a 'joint' evaluation of the course using a discussion/survey format. The survey summarized the instructor's reflections on the course noted throughout the term and asked students to rate the application or feasibility of those comments to incorporate into future courses to enhance overall effectiveness and efficiency. Instead of just administering the survey, however, the instructor introduced each question to describe the rationale for the reflection/survey item to each student. Thus, the instructor would introduce a topic for evaluation, allow students to discuss it openly, and finally allow them to record their responses on the survey form. Then, the next topic would be introduced and evaluated. A copy of the student evaluation feedback form is attached. Additionally, the typical end of course survey at the department and college level was administered online outside of this classroom session.

Results

Both quantitative and qualitative survey data (see attached – PL476 After Action Review) illustrate how student attitudes and suggestions compared to instructor reflections on numerous aspects of the course such as scheduling, learning activities, and course redesign. Students also provided valuable feedback to enhance the learning process. The two sections from the first year of the study provided feedback towards the initial revisions made to the course before the term and also provided excellent suggestions for subsequent iterations of the course, which were then incorporated into the second year evaluation.

The number of requirements and the frequency in which they were required during the first year of evaluation, led to the course being rated by students as significantly more demanding and rigorous than other courses in the department and at the college by. Students also made comments such as "I needed more time to practice or apply the material I was learning throughout the semester" and "too many graded events" on the end of course survey.

Interestingly, the students, who worked individually on their design project in the first year, strongly opposed the instructor's suggestion for collaborative work in the future. Instructors then revised the undergraduate course with a 2-person collaborative project instead of an individual project. Students strongly supported this change after the course, contrary to previous students' recommendations to maintain individual projects.

Results of the typical university end of course survey revealed that students valued the learning, instructional activities, and delivery in this course more than other courses at the college and other courses in the academic major of the course. Students also provided likes-dislikes in the course structure and delivery as proposed by Hampton & Reiser (2001). Keller's ARCS model (1987a, 1987b) was noted as one of the most valuable applications of the course.

The instructors categorized the responses from the students by Reiser & Dick's six instructional activities: M = motivating students; O = Inform Students of objectives; H = help recall prerequisites; I = present information; P = practice and feedback; S = summarize lesson. The majority of the responses expressing dislike from the first year (2001) centered on information presentation and practice and feedback. Some examples of those comments include:

(I) "Two periods of [classroom observation] briefs were repetitive"

(P) "Drop the final PPT [powerpoint] presentations on the classroom observations"

(P) "There were a lot of graded assignments, I think that cutting down on a few of these would allow us to spend a little more time on each PPO" [Product Performance Objectives]

Based on comments such as these and the statistical data taken from the student evaluation feedback form and end of course surveys, the instructors reduced the number of graded assessments from thirteen individual items to seven individual items and three group projects and changed one presentation to a roundtable discussion. By altering the requirements, the instructors increased the average amount of time students were given to complete assignments by 1.2 lessons and reduced the instructor grading load by 35%.

The results of the end of course survey indicate an increase in student satisfaction for many areas of the course. On average, this course was rated higher in the areas of student motivation, teacher effectiveness, and perceived learning than other courses in the department and at the Academy (see figure 1 – End of Course Student Survey Results). Sample comments on the end of course survey the second year include:

(I) "Definitely [sustain one lesson of roundtable discussion instead of classroom observation briefs] – gets the ideas flowing – maybe even do it sooner in the course."

(P) "I agree that each member in the 2-person team should start out with the same topic instead of two different topics."

(P) "I liked having the assignments due throughout the semester to help with preparation, so it is not saved to the last minute."

Selected End-of-Course Student Survey Questions	USMA or Dept.		Course '01		Course '02	
	M	SD	M	SD	M	SD
1. Inst. used effective techniques for leaning both in class and for out-of-class assignments	**4.22**	.87	**4.43**	.73	**4.56**	.5
2. My motivation to learn has increased b/c of course	**3.93**	1.03	**4.24**	1.02	**4.33**	.47
3. In this course my critical thinking ability increased.	**4.0**	.95	**4.0**	.93	**4.33**	.67
4. I understood the learning outcomes desired for this course.	**4.09**	.87	**4.24**	.81	**4.44**	.68
5. The activities in this course helped me learn more about myself as a person.	**4.14**	.92	N/A		**4.56***	.5
6. This course has increased my flexibility & adaptability for future learning & problem solving.	**4.07**	.85	N/A		**4.33**	.82
7. This course has inspired & challenged me to learn more about this topic on my wn.	**3.84**	1.06	N/A		**4.22**	.92
8. Compared to other courses at West Point, this course was more demanding.	**3.07**	.93	**3.86***	.71	**3.78**	.92

Figure 1
Scale range: 1= Strongly disagree to 5 = Strongly agree * $p < .05$

Discussion

The systematic process used to enhance the effectiveness and efficiency of the course in this study could be replicated in any academic course. Techniques discussed at AECT 2001 (Reiser & Hampton, 2001) for systematically evaluating the instructional practices of a course were quite applicable in this study. Ultimately, it enhanced what McKeachie (1997) said about effective teaching: "Good teaching involves building bridges between what is in your [the teacher's] head and what is in the students' heads" (p.1223).

References

Ingram, K. (2001). *The effects of reflective thinking training on TAs' reflective thinking, use of instructional activities, instructional effectiveness, motivation to teach and their students' attitudes toward instruction.* (Doctoral Dissertation, Florida State University). *Dissertation Abstracts International, 62/2,* 486.

Keller, J. M. (1987a). Strategies for stimulating the motivation to learn. *Performance & Instruction,* 26(8), 1-7.

Keller, J. M. (1987b). The systematic process of motivational design. *Performance & Instruction,* 26(9), 1-8.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, *52(11),* 1218-1225.

Reiser, R. A. & Hampton, S. E. (November, 2001). Practicing What We Preach: Effective Teaching through the Use of Instructional Design Principles. Association for Educational Communications and Technology (AECT) International Conference (Presidential Session), Atlanta, GA, November.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* San Francisco: Jossey-Bass.

PL476 After Action Review

Please read each of my considerations and circle your level of agreement, from strongly disagree (1) to strongly agree (5), beside the revision. If you have more detailed comments or suggestions, please make notes in the question box or below the table.

Suggested Revisions to PL476	1=strong disagree 5=strong agree	
1. Changing Ed Pscych text (Psychology of Teaching), which is geared toward K-12 grade, to Driscoll's book on The Psychology of Learning for Instruction (we used excerpts for Gagne and Keller's model).	1 2 3 4 5 (2.23)	(3.86)
2. Add more time to prep PPO3 (Instructional Analysis with all those boxes☺)	1 2 3 4 5 (3.71)	
3. Do the planning phase of this course in a 2-person team. e.g. ID your goal, instructional analysis, objectives/tests, and instructional strategy as a team (PPO1-6). Then, rehearse, deliver and evaluate the instruction as an individual (1:1, Teach, Small Grp eval).	1 2 3 4 5 (1.81)	(4.13)
4. Incorporate 75-100 'instructor points' for the course. Factors that instructor would consider are: classroom participation, effectiveness/effort of the mini-practice teach you do early in the course (see item 5 below), timeliness of graded assignments, and possibly a small portion of teammate input (e.g. level of support in 2-person team).	1 2 3 4 5 (2.90)	(3.67)
*2002 Only – Keep the Daily Study Questions (DSQs) to keep you reading the assignments to maintain your participation in class.	1 2 3 4 5	(3.73)
5. Assign students to do a 5-minute attention getter, give example or practical exercise for a classes early in the term to get used to 'teaching' in front of class BEFORE you have to teach your final. This would be P/F grade, but could also influence the 'instructor grade' points.	1 2 3 4 5 (3.48)	
6. Drop PPO 8 (revised & highlighted strategy)	1 2 3 4 5 (2.85)	
7. Drop 1 of the 4 lessons on Obj/assessment (I think we overkill this area)	1 2 3 4 5 (4.43)	
8. Use Plebes as students for your final classes. You would still attend several-but not all. 8.5 Have peers in rehearsal	1 2 3 4 5 (4.25) (3.29)	(3.93)
9. Drop 2 days of briefs on classroom observations (Lsn 38/39) and make it one lesson of roundtable discussion, sharing strengths, weaknesses, etc. of observed instructors. *(**note 2002 question revised to say "Sustain roundtable discussion…")*	1 2 3 4 5 (4.42)	(4.53)
10. Instead of talking these briefs, submit a short 2 page summary of observations on the handout I gave you and stapling the checklist to it.	1 2 3 4 5 (3.52)	
11. Move class observation discussions to the class immediately following our class on instructional delivery. This would require you to go out and observe/evaluate another instructor around lesson 22 BEFORE you have to teach your final class. Will also give you more time to prep your lesson, rehearse etc. for those who go earlier.	1 2 3 4 5 (4.33)	
12. Adding another class on Instructional delivery where you actually see examples of prior students (YOU☺) demonstrating delivery techniques. This would then give you 3 lessons on instructional delivery (vs. 1 this term). 1st =class like this term, 2nd = Video examples of cadets 'showing' you delivery in action, 3rd=roundtable discussion of your observations of other instructors you observed	1 2 3 4 5 (3.62)	
13. Covering @ chapter of Dick & Carey w/ 2 lessons: 1 lsn to review the material from the chapter, see examples etc. The next lsn to work on YOUR course project in your 2-person teams w/ instructor available & summarizing key points/examples at the end of class. In essence, you get started on your project homework in a working class session.	1 2 3 4 5 (3.81)	(4.67)
	2001	**2002**

simNewton as an Aid to Conceptual Learning of Newtonian Motion

Leo Hirner
Metropolitan Community College and University of Missouri

Deanna Poudel
Longview Community College and Iowa State University

Abstract

The growth of online courses has brought with it questions of how to manage laboratory instruction. One potential solution to the lab problem is the use of interactive simulations that not only recreate a physical system, but also stimulate the building of conceptual models necessary for an introductory student to develop an understanding of physics. This paper examines the effectiveness of the simNewton interactive modeling tool for replacing traditional labs, its effectiveness in building students' conceptual understanding, and whether introducing the lab earlier or later impacts student learning.

Introduction

The growth of web-based courses has generated a number of issues for lab science instructors. How do students satisfactorily complete the laboratory component of the course without coming to campus? Are home labs the equivalent of on-campus labs? Can simulations replace some or all of traditional hands-on lab? While these questions remain to be answered, the use of interactive, online simulations does appear to be one potential avenue for meeting lab instruction needs and improving student understanding and retention of physics concepts.

The concept of "learning by doing" is at the heart of both the behaviorists' and constructivists' support of lab instruction. Leonard noted that quality lab instruction contained at least three characteristics: student engagement in scientific inquiry, ability of students to manipulate the components, and a student experience designed to build understanding of the theoretical model (1989). The weakness of the traditional lab design is often in the prescribed nature of the lab exercise, which rarely includes any one of Leonard's criteria, especially building any connection between theory and application. The potential of simulations for assisting students to build mental models, transfer concepts, and apply theoretical knowledge to the physical world has been shown in a variety of studies (Lunetta & Hofstein, 1981; Choi & Genaro, 1987; Thomas & Hooper, 1991).

The recognition that student preconceptions form a barrier to conceptual understanding of physics principles as well as problem solving methods has revealed the need to re-evaluate the basic paradigms of physics education. Over the last 20 years, this has resulted in the development of a number of new tools and methods for instruction. Cognitive research on student learning in traditional physics education (Larkin, 1981; Clement, 1982; McDermott, 1984; Halloun & Hestenes, 1985; Mestre & Touger, 1989, Hestenes, et al., 1992) has generated a taxonomy of the preconceived misconceptions about the physical world, how these misconceptions interfere with student learning and strategies for overcoming these pre- and misconceptions within the topic of mechanics, especially kinematics. Physical models and definitions of motion are clearly not congruent with the preconceptions of motion that students generally have when entering a physics course. Preconceptions that interfere with student learning include explicit definitions of velocity (such as distance divided by time), no recognition of the difference between velocity and acceleration, no recognition that changes in direction imply acceleration, and that when an object reaches the top of its trajectory (momentarily coming to rest in the vertical direction) that no acceleration exists.

Several people have studied the differences between novice and expert problem solvers, especially the mental models each has of a given physical system, and persistence of the novices' naïve theories. One of the advantages held by experts is the "index" of experiences that assists the expert to readily identify the important information associated with a given problem system (Shank, Berman & Macpherson, 1999). How to get the novice student to recognize their misconceptions, tear down their mental constructs, and then build new mental models is the challenge.

Traditional physics instruction, supported by behaviorist theory, relies upon the teacher-centered method of lecture and example. This has proven to be ineffective at attending to the persistent preconceptions students have before entering a college level science course. On the other hand, constructivists' fundamental assumption is that an individual's knowledge is a consequence of their experiences and their construction of these experiences into body of knowledge. One of Mestre's conclusions is that it is important to provide students experiences that call into question their naïve theories. He advocates active participation of the learner in confronting and discussing contradictory explanation or predictions of physical phenomena (1991).

Research beginning with Halloun and Hestenes in 1984 revealed to physics educators that the didactic methods by which physics has historically been taught are generally effective for those that learn much like the traditional faculty. Over the next few years these studies led to new perspectives on teaching physics, especially a focus on process not content and a need to better understand the cognitive processes employed by students began to emerge (Mestre & Touger, 1989; Redish, 1994; Hestenes, 1998; Redish, 1998). These studies even led to system projects that promoted these teaching methods through professional groups such as the NSF funded Two-Year Colleges in the 21st Century initiative of the American Association of Physics Teachers begun in 1991 (http://www.tycphysics.org).

Radically different approaches based on constructivist learning theory, although effective, are not practicable within the context and time constraints of current educational programs. Instruction that focuses on building causal models bridges the extremes. Teaching via causal models begins with simple representations of the system, and then directed activities guide students in applying causal reasoning to illustrate the mechanisms or models of the simple system at an intermediate level (White, 1993).

Instructional design for building causal models begins with the simple introduction of the concept via a "Source Model." The Source Model serves to introduce the concept through the most basic representations that are easily understood by or familiar to the students. Students require time to work with and master manipulating the Source Model to set up introduction of the "Derived Model." The Derived Model bridges the gap from simple to intermediate level conceptualization, and the Derived Model must be connected conceptually to the Source Model. The Derived Model must build upon the basic concept(s) of the Source Model, thus creating conceptual linkages between the two models and promoting transferability of the concept to generalized examples (Fredericksen, White & Gutwill, 1999).

Importance of Simulations

"Instructional computer simulations" are computer programs based upon mathematical models based upon either a physical model or theoretical prediction of the natural world that may be manipulated by the student user, often without the distractors present in most traditional representations (Weller, 1996). Thomas & Hooper noted that interactive programs allow manipulation from the introductory state through a number of intermediate steps to the final, or goal, state (1991).

Thomas & Hooper identified four applications for simulations in instruction:

- Experiencing simulations that prepare students for concepts to be examined in a traditional manner.
- Informing simulations used to augment or replace textbook instruction.
- Reinforcing simulations are designed to stay within the instructional context and reinforce learning.
- Integrating simulations bring together a set of separate facts or concepts and combine these into a global model of the systems studied.

Microworlds go beyond the minimally interactive simulations. These microworlds are based on an underlying set of mathematical models consistent with the global set of observable phenomena. The environment not only allows to interact with the simulation, but students are able to create their own simulations within the constraints of the global model (Weller, 1996).

Jonassen points out that interactive simulations are able to meet two of the three elements of student-centered learning environments (the other element being context). In addition to the simulation space, the interactivity of the simulations provides students the opportunity to manipulate the system beyond the traditional constraints (2000).

The use of microworlds, such as *Interactive Physics*™*,* for assisting students in building conceptual models is supported by several theoretical perspectives (Jonassen, Peck & Wilson, 1999; Land & Hannafin, 2000). While simNewton does not have the range of capabilities of *Interactive Physics*™*,* it does allow for the creation of web-based, manipulable simulations that can be used to develop and reinforce student's creation of causal models.

Simulations provide a potential solution for overcoming existing student misconceptions. Experiencing simulations that directly challenge misconceptions were found to influence students in recognizing the existence of problems in their personal "models" (Posner & Strike, 1989). Subsequent research by Flick (1990), Weller (1995), Henessey et al (1995 a, 1995b) and Gorsky and Finegold (1994) supported the use of interactive simulations as a tool for overcoming erroneous student preconceptions.

simNewton overview

simNewton is an interactive learning environment built on ThinkerTools, a simulation software package developed by White, Frederiksen, et al (1993). simNewton was created by Han Chin Liu (2002) to be a web-based implementation of ThinkerTools. In this environment, students see a "dot" representing an object moving, at first, without friction or gravity. They are asked to make the dot complete a task such as landing on an "x." The students can use the keyboard to apply a horizontal or vertical impulse, which is a hit or a force acting for a very short duration, to change the motion of the dot. The dot's trajectory can optionally be graphed as a dot print (rather than a foot print) path and the motion can be restrained with walls drawn to limit the motion of the dot.

The simNewton lessons used follow White, et al's intermediate causal model. Each segment begins with simple representations of motion in one and two dimensions, and then builds to more complex representations tied to the initial simple representation of the concept. Activities progress from very simple tasks such as landing on the dot at a certain velocity, to landing a spaceship or guiding a massive object around a maze. Friction and gravity can be added in after the students have experience without them. Within the activities, the students make predictions of how the even will go, and then are asked to reflect on how the simulation matched their predictions. The student-controlled nature of the environment coupled with the student's ability to test various situations under constraints places it in the category of a microworld (Jonassen et al., 1999). The figure below shows some examples of the student interface for simNewton.

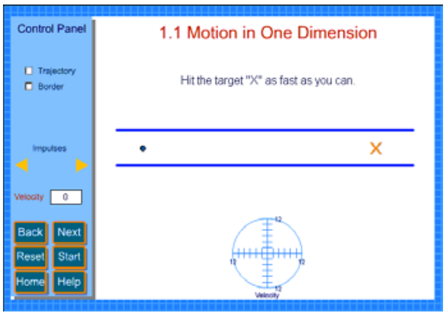


Figure 1: One of the simpler tasks

Figure 2: A more complicated task

The original ThinkerTools was produced in collaboration between the University of California, Berkeley, and the Educational Testing Service. It has developed into a comprehensive middle school curriculum for learning Newtonian mechanics and includes instructor materials, assessments and other materials for implementation in the classroom. The curriculum includes real activities intermingled with the simulations, so that the students can bridge their understanding between the real and virtual worlds. For example, before beginning the first simulation activity, the students are given a real croquet mallet and a ball. They practice hitting the ball so that they can deliver a unit of impulse to the ball consistently, measured by a sonic ranger type device or by measuring the distance the ball moves in the presence of friction such as on carpet. The goal is that the ball changes its velocity by the same amount for each hit. After achieving this, the students are asked to complete simple tasks similar to what they will do in the simulation, such as making the ball hit the target at a certain velocity. With this experience, then, the students interact with the simulations, and can transfer the real motion of the ball to the virtual motion of the dot. The predictive aspects of the scientific process are developed in depth in the curriculum. Students can make their own hypotheses and design and conduct virtual experiments to test them. Students draw their own conclusions base don't the evidence they see in the experiments.

Although ThinkerTools has been shown to be effective as described above (Whit, Fredericksen, 1999), Liu looked at whether a subset of the simulation activities of this type could be used independently and effectively. Liu created a package of these activities in simNewton, updating the interface to be web-based (2002). Liu's study tested simNewton with a group of middle school students and found these activities to be useful experiences in learning the basic principles of Newtonian motion.

In contrast to these previous studies, the present study observes the simNewton activities as they are used as part of the required curriculum for an online community college course in physical science, rather than with volunteer groups of middle school students. Research on characteristics of online community colleges students indicates that this population is more self-reliant and motivated than traditional college and high school students (Kennedy, 2000; Halsne, 2001). There is interest in seeing how this population responds to the activities and the potential for developing causal models through the simulations.

Previous studies have shown that the order of activities in the unit affects student outcomes. For instance, Brant, Hooper and Sugrue (1991) found that using simulations before didactic instruction improved student scores. Studies conducted specifically using ThinkerTools gave similar results. This is supported by several learning theories that suggest that students must have an experience base from which to draw before being able to use symbolic models of the phenomena. (Andre et al., 1996) In the context of online classes, this offers a challenge to instructors as to how to guide the order of activities. Students are able to choose the order or their study activities to a great extent and in spite of encouragement from the instructor may complete the components of a unit in a different order. Course management software allows some control of the order that students can access materials, but this varies from platform to platform. The current study uses such controls to direct the order in which students complete the lab activity and text-based activities in an online course.

The Experiment

This investigation took place at a medium-sized community college located in a large mid-western, metropolitan area. Students were enrolled in an online physical science course with no pre-requisite requirements. The instructional goals of this course focus on building a conceptual understanding so students are not assessed on traditional problem solving ability.

Student access to the course is made through the WebCT^{TM} course management system, and the designers have taken advantage of the conditional function in WebCT^{TM}'s testing and surveying tools to ensure that the treatment groups followed their instructional paths. Before beginning instruction, the students in the treatment groups completed a pre-test assessing their preconceptions. Submission of the pre-test was the trigger that released the instructional components of the course. The first treatment path has the students complete a review of the lecture notes, homework, and discussion activities before completing survey assessing their epistemological beliefs. Although this data is not analyzed here, completion of the survey was used to trigger the opening of the lab exercise link. The second treatment reversed the activities before and after the epistemological belief survey. The control format released all instructional elements simultaneously, and it used a traditional lab exercise using cars rolling down a slight incline to measure acceleration.

At the conclusion of the instructional module all students completed a post-test to assess student learning. The pre-test used questions from the Halloun & Hestenes assessment (1984), the student attitude survey is based upon a tool designed by Schommer (1993), and the post-test included content questions from the original course materials and the Force Concept Inventory (Hestenes et al., 1992).

Hypotheses-

- Students that undergo either of the treatment paths will perform better on the conceptual assessments than the control group.
- The two treatment paths will not demonstrate a significant difference in their performance on the assessment.

The pre-test, post-test, survey instruments, the lab and simNewton activities may all be found at http://www.kcmetro.cc.mo.us/distance/hirner/simnewton/simnewton_index.htm.

Data Analysis

The first two experimental groups (simNewton before content, SBC, and simNewton after content, SAC) completed the treatment during the spring semester while the control group, CON, completed the traditional lab with the content during the fall.

Pre-test

All students completed a pre-test consisting of ten questions directly addressing fundamental motion from the Force Concept Inventory prior to the release of content and lab exercises (possible using the conditional controls within the course management system). An initial test for Homogeneity using Levene's test, $F(5,55)=..378$, $p > 0.05$, and the kertosis, 1.246 and skew, -0.547, of the data illustrated that the data set fit the general assumption.

No significant difference was found between the three groups (CON, M = 5.00, SD = 1.612; SBC, M = 5.33, SD = 1.455; SAC, M = 5.59, SD = 2.282), $F(2,50) = 0.021$, $p > 0.05$. Analysis for temporal influences were not significant, $F(1,60) < 0.001$, $p > 0.05$, eliminating concerns associated with the delay between course offerings. Gender analysis was not possible as the subjects in all cases were skewed towards the female (CON, N = 21, male = 2; SBC, N = 18, male = 2; SAC, N = 22, male = 2).

A test for correlation was completed between the scores of the pre-test and that of the post-test (Module 2 Quiz) and very little correlation was found. When the same test was completed on the reduced set of questions (explained below) the result was similar ($R^2 = 0.0316$). The rest of the analysis was completed without correlating the Module 2 Quiz scores with the pre-test scores.

Module 2 Quiz

At the conclusion of the instructional unit the students completed a twenty-question quiz composed of eleven kinematics questions, from the Hestene's test, and nine questions from a pool assessing additional content. Testing for homogeneity using Levene's test, $F(5,55)=0.512$, $p > 0.05$, and the kertosis, 0.007 and skew, 0.307, of the data illustrated that the data set fit the general assumption.

No significant difference was found between the three groups (CON, M = 10.95, SD = 3.154; SBC, M = 12.50, SD = 3.400; SAC, M = 11.14, SD = 3.044), $F(2,50) = 1.516$, $p > 0.05$. Analysis for temporal influences were not significant, $F(1,60) < 0.001$, $p > 0.05$, eliminating concerns associated with the delay between course offerings, and, as with the pre-test, the gender distribution did not lend itself to analysis

Hestenes Questions

A more detailed analysis of subject performance on the impact of the simNewton activities was comprised of the eleven questions from the Hestene's set. Reliability tests were run on each item and two were found to be questionable, reducing the reliability coefficient alpha. Removing these two questions produced increased Alpha from 0.703 to 0.780. Testing for homogeneity on the reduced set of nine questions using Levene's test, $F(2,58)=0.969$, $p > .05$, the kertosis, -0.745 and skew, 0.004, of the data confirmed that the data fit the general assumption.

Group	Mean Score	SD	Lab Activity	
CON	4.57	2.675	Traditional Acceleration Lab (written)	F(2.61) = 2.304, p > 0.05
SBC	6.17	2.007	SimNewton lab before content was given	
SAC	4.95	2.400	SimNewton simulation after content was given	

No significant difference was found between the three groups (CON, M = 4.57, SD = 2.675; SBC, M = 6.17, SD = 2.007; SAC, M = 4.95, SD = 2.400), $F(2,61) = 2.304$, $p > .05$. Analysis for temporal influences were not significant, $F(1,60) < .001$, $p > .05$, eliminating concerns associated with the delay between course offerings, and, as with the pre-test, the gender distribution did not lend itself to analysis.

Effect Size and Power

In all three analyses the effect size increased from low in the pre-test, $\eta^2 < .01$, to medium in the Module 2 Quiz, $\eta^2 = 0.027$, to high in the Hestene's questions, $\eta^2 = 0.074$. While the sample sizes were well within those determined using Pearson-Hartley tables (for a power of 0.80 a minimum sample size of 15 for each treatment was estimated), the resulting power determined by SPSS for each case were all < 0.50.

Discussion

The no significant difference results in both the Module quiz and reinforced by specifically examining only the results on the questions from the Hestene's test imply that the web-based simulations using simNewton were at least as effective as the traditional lab exercise in supporting students' development of a conceptual understanding of kinematics. Further, the results do not support the findings of Brant, et al. that introducing the lab simulation prior to content improved the development of the student's conceptual framework.

The results of this study are limited by two the subjects being of primarily one gender and community college students. The results may only be generalized to female students due to the large imbalance in gender (approximately 90% in each treatment group). This is further limited as the students self-selected the section through the institutional enrollment process. Greater coordination with the hosting institution is necessary to generate a randomly assigned sample with a truer gender mix.

In all cases the subjects were non-traditional community college students, though this distinction is beginning to blur. Enrollment patterns at the institution where this experiment was performed are showing a trend of traditional students augmenting their schedules with online courses. In 1998 there were less than 500 online students taking an average of 1.2 courses online. Enrollment in the fall of 2001 found more than 2000 students taking 1.5 courses online with a total enrollment of three courses. Further, the average age of students in 1998 exceeded 30 years, and in 2001 the student population was almost evenly split between the over 30 population and a traditional age (18-22) group of students.

There is a need to continue to study the methods used in online instruction at community colleges with recent national enrollment data noting that ninety percent of two-year institutions offered online courses as of the 2000-01 academic year and account for forty-eight percent of the enrollment nationwide that year (Watts & Lewis, 2003). Coupling this with the self-selecting characteristics of online students noted by Kennedy (2000), and the self-efficacy and motivation observed by Halse (2001), and the need to learn more about both the learning styles and motivational factors of online students at the community colleges is clear.

While the study demonstrated that the simulation laboratory exercise was at least as effective as the traditional physical lab exercise, there was hope that the simulation would prove more effective. The conflicting power results indicate that a more extensive sample may provide an answer as to whether the simulations are more effective or not. A study needs to be carefully designed to gather a balanced population in terms of gender, and it would be of interest to see if the age of the participants is a factor in the outcomes or effectiveness of the simulations. There is anecdotal evidence that younger students of the "video game" generation are more likely to respond and interact with simulation tools. A more extensive study accounting for these additional factors should extend the educational community's understanding of the impact that newer, web-based, interactive tools may play in developing conceptual understanding of physical systems.

References

Andre, T (1996) Mission Newton! and Thinker Tools: Using prior simulations to promote learning about motion. Unpublished manuscript, Iowa State University.

Brant, G., Hooper, E., & Sugrue, B. (1991). Which comes first the simulation or the lecture? Journal of Educational Computing Research, 7(4), 469-481.

Choi, B.-S., & Gennaro, E. (1987). The effectiveness of using computer simulated experiments on junior high students' understanding of the volume displacement concept. Journal of Research in Science Teaching, 24(6), 539-552.

Clement, J. (1982). Students' preconceptions in introductory mechanics. American Journal of Physics, 50(1), 66-71.

Frederiksen, J. R., & White, B. Y. (1998). Teaching and learning generic modeling and reasoning skills. Interactive Learning Environments, 5, 33-51.

Frederiksen, J. R., White, B. Y., & Gutwill, J. (1999). Dynamic mental models in learning science: the importance of constructing derivational linkages among models. Journal of Research in Science Teaching.

Halloun, I. A., & Hestenes, D. (1985). Common sense concepts about physics. American Journal of Physics, 53(11).

Halsne, A.M. (2001). Online versus traditionally-delivered instruction: A descriptive study of learner characteristics in a community college setting. Doctoral dissertation, Loyola University Chicago.

Heller, J.I., & Reif, F (1984), Cognition and Instruction, 1**,** 177.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. The Physics Teacher, 30, 141-158.

Hestenes, D. (1998). Who needs physics education research? American Journal of Physics, 66(6), pp 465-7.

Jonassen, D. H., Peck, K. L., & Wilson, B. G. (1999). Learning with Technology: A Constructivists Perspective. Upper Saddle River, NJ: Merrill, Prentice Hall.

Jonassen, D. H. (2000). Revisiting activity theory as a framework for designing student-centered learning environments. In D. H. Jonassen & S. M. Land (Eds.), Theoretical Foundations of Learning Environments (pp. 89 - 121). Mahwah, NJ: Lawrence Erlbaum Associates.

Kennedy, C.A. (2000). What influences student learning in an online course? ED466238

Land, S. M., & Hannafin, M. J. (2000). Student-centered learning environments. In D. H. Jonassen & S. M. Land (Eds.), Theoretical Foundations of Learning Environments (pp. 1 - 19). Mahwah, NJ: Lawrence Erlbaum Associates.

Larkin, J. (1980). Cognition of Learning Physics. American Journal of Physics, 49(6), 534-541.

Leonard, W. H. (1989). Ten years of research on investigative laboratory instruction strategies. Journal of College Science Teaching, 18(5), 304 - 306.

Liu, Han-Chin (2002). Investigating the use of ThinkerTools to promote learning of Newton's laws of motion among middle school students. Doctoral Dissertation Iowa State University.

Lunetta, V. N., & Hofstein, A. (1981). Simulations in science education. Science Education, 3, 243-252.

McDermott, L. C. (1984). Research on conceptual understanding in mechanics. Physics Today, 24-32.

McDermott, L. C., & Redish, E. F. (1999). Resource Letter on Physics Education Research.: United States Department of Education.

Mestre, J., & Touger, J. (1989). Cognitive research-what's in it for physics teachers. The Physics Teacher, 447-456.

Mestre, J. (1991). Learning and instruction in pre-college physical science. Physics Today, 56-62.

Posner, G. J., & Strike, K. A. (1989). The conceptual ecology of physics learning. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Redish, E. F. (1994). The implications of cognitive studies for teaching physics. American Journal of Physics, 62(6), 796-803.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1997). On the effectiveness of active-engagement microcomputer-based laboratories. American Journal of Physics, 65, 45-54.

Redish, E. F. (1998). Millikan Award Lecture: Building a science of teaching physics. American Journal of Physics, 67, 562-573.

Schommer, M. (1993). Comparisons of beliefs about the nature of knowledge and learning among postsecondary students., University of Illinois, Urbana-Champagne, IL.

Schank, R. C., Berman, T. R., & Macpherson, K. A. (1999). Learning by doing. In C. M. Reigeluth (Ed.), Instructional-Design Theories and Models: a New Paradigm of Instructional Theory (Vol. II, pp. 161 - 182). Mahwah, NJ: Lawrence Erlbaum Associates.

Thomas, R., & Hooper, E. (1991). Simulations: an opportunity we are missing. Journal of Research on Computing in Education, 23(4), 497 - 514.

Trowbridge, D. E., & McDermott, L. C. (1980). Investigation of student understanding of the concept of velocity in one dimension. American Journal of Physics, 48(12), 1020-1028.

Watts, T. & Lewis, L. (2003). Distance education at degree granting postsecondary institutions: 2000-2001. NCES 2003-017. U.S. Department of Education, National Center for Educational Statistics, Washington, D.C.

Weller, H. G. (1996). Assessing the impact of computer-based learning in science. Journal of Research on Computing in Education, 28(4), 461-486.

White, B. Y., Frederiksen, J. R., & Spoehr, K. T. (1993). Conceptual models for understanding the behavior of electrical circuits. In M. Caillot (Ed.), Learning Electricity and Electronics with Advanced Educational Technology. New York, NY: Springer Verlag.

White, B. Y. (1993). Intermediate Causal Models: A missing link for successful science education. In R. Glaser (Ed.), Advances in Instructional Psychology (Vol. 4, pp. 177-251). Hillsdale, NJ: Lawrence Erlbaum Associates.

White, B. Y., Shimoda, T. A., & Frederiksen, J. R. (1999). Enabling students to construct theories of collaborative inquiry and reflective learning: Computer support for meta-cognitive development. International Journal of Artificial Intelligence in Education, 10(2).

White, B. Y., & Frederiksen, J. R. (in press). Technological tools and instructional approaches for making scientific inquiry accessible to all. In M. Jacobsen & R. Kozma (Eds.), Learning the Sciences of the 21st century: Theory, Research, and the Design of Advanced Technology Learning Environments. Mahwah, NJ: Erlbaum.

White, B. Y., & Scwarz, C. V. (in press). Alternative approaches to using modeling and simulation tools for teaching science. In N. Roberts & W. Feurzeig & B. Hunter (Eds.), Computer Modeling and Simulation in Science Education. New York, NY: Springer-Verlag.

Student Experiences in and Perceptions of Field Experiences: Virtual and Traditional

Emily Hixon
Hyo-Jeong So
Indiana University

Introduction

Each semester, multiple sections of Introduction to Educational Psychology are offered at Indiana University's School of Education. This course has an accompanying early field experience that requires students to be placed in local schools for at least 20 hours over the course of the semester. Since many classes include such a field experience and all pre-service teachers must also be placed for student teaching, the local schools are becoming overwhelmed with the number of Indiana University students in their buildings. To alleviate some of this burden, school officials are looking at alternatives for field experiences. One such option is the virtual field experience that utilizes the video-based cases included in the Inquiry Learning Forum (ILF).

Before the School of Education adopts such a program on a larger scale, it was important to understand how the experience of the students completing the virtual field experience compares with that of the students completing the traditional field experience. This study was designed to explore students' experiences in and perceptions of the different settings and attempted to understand some of the mechanisms underlying student views.

Literature Review

Field experiences (including early pre-student teaching and student teaching experiences) are a foundation in preservice teacher education programs today. Although much research has reinforced the value of field experiences as necessary components of teacher education programs (e.g., Aikens & Day, 1999; Li & Zhang, 2000; Metcalf & Kahlich, 1996; Paese, 1985), concerns and limitations have also been expressed related to the quality and impact of field experiences. Of major concern to teacher educators is the fact that prospective teachers may not be cognitively prepared to benefit from an early field experience (Aiken & Day, 1999; Feiman-Nemser & Buchmann, 1985; Goodman, 1986; Hannah, 1995). Students who are in the early stages of their teacher education program may not have the knowledge and "tools" they need to be able to interpret and critically analyze the classroom environment and events.

Another major concern is that students are not taking an active role in developing, processing, and reflecting on their field experiences (Feiman-Nemser & Buchmann, 1985; Goodman, 1986; Johnston, 1994). On a more practical note, with the impending teacher shortage resulting in growing preservice education programs, schools of education are being forced to place students at sites that are at increasing distances from campus

In an effort to address concerns with traditional field experiences, educators are exploring the role that various technologies may play in supplementing and perhaps replacing traditional field experiences. Some of the technologies being considered include: videodisks (Brooks & Kopp, 1991; Chaney-Cullen & Duffy, 1999; McIntyre & Pape, 1993; Lambdin, Duffy, & Moore, 1996), videotapes (Atkins, 1998), two-way interactive television/video conferencing (Garrett & Dundt, 1998; Gruenhagen, McCracken, & True, 1999; McDevitt, 1996), and online asynchronous conferencing (Bonk, Malikowski, Angeli, & East, 1998; Mason, 2000; Moffett, 2001). Several specific benefits of technology's use in relation to field experiences include:

- Reducing costs
- Creation of shared experiences and quality control
- Promoting reflectivity
- Preparing students cognitively
- Exposure to various teaching styles and strategies
- Connection to class content
- Allowing for review of events
- Learning about technology

When technology is used to completely replace a field experience, new issues and concerns must be addressed. Atkins (1998) discussed a situation where the field experience component was cut for an elementary mathematics methods course, leading instructors to turn to technology to help students "construct a vision of effective mathematics teaching" (p. 99). Specifically, students viewed, engaged in discussions, and completed activities on several videotapes related to effective practices in mathematics education. At the end of the semester, over half (58%) of the participants stated that the lack of a field experience did not affect their understanding of effective mathematics teaching. However, nearly a quarter of the students (24%) did feel that the lack of a field experience affected their understanding, citing their inability to test out the ideas and see them being used with children as the main reasons for this concern. The authors conclude that "many of the respondents associated the lack of a field experience with a lack of opportunity" (p. 95).

Lambdin, Duffy, and Moore (1996) also found that students had similar concerns about the Strategic Teaching Framework as one component of their field experience. Strategic Teaching Framework (STF) is an interactive videodisk information system designed to help preservice teachers develop teaching strategies through a constructivist learning environment. While most students appreciated the use of STF, almost all believed that hands-on time in the classroom was even more valuable to them, and STF could not be substitute for actual classroom experience.

These are legitimate concerns that must be considered when technology is used to replace a traditional field experience. While the technology may offer some advantages (as suggested above), it may also present new challenges. The benefits and limitations of using technology related to field experiences must be weighed carefully.

Current Study

Due to concerns related to the number of students requiring field placements, Indiana University's School of Education is exploring possible alternatives for one of its early field experience requirements which accompanied the Introduction to Educational Psychology course. The purpose of the current study was to explore students' perceptions of and experiences in the virtual field experience as compared to students completing the traditional field experience. Specific questions being addressed include:

- Do students completing the virtual field experience recognize and take advantage of opportunities that the traditional field experience may not offer (e.g., more time for reflection, observation of more teaching styles and student characteristics, more opportunities to see class concepts "in action")?
- Do students completing the virtual field experience view it as authentic? Do they differ from those completing the traditional field experience in their ability to "get a feel" for being in the classroom?
- Do students completing the virtual and traditional field experiences differ in terms of their satisfaction with and value gained from their field experience? What strengths and weaknesses do students perceive for their specific field experience format?

Method

Participants

Students from four sections of an Introduction to Educational Psychology course were invited to participate in this study. Two of these sections were traditional "live" offerings of the course, while the other two sections were web-based classes offered completely online. Students in one of the web-based sections (N=20) participated in the virtual field experience; students in the other web based class (N=22) and in the traditional sections (N=64) completed a traditional field experience.

Table 1 *Sections of Introduction to Educational Psychology*

	Sections 1 & 2	Section 3	Section 4
Course	Live	Web-based	Web-based
Field Experience	Traditional	Traditional	Virtual
N	64	22	20
n	32	10	8

The Field Experiences

Traditional Field Experience. Students completing the traditional field experience were placed in a local classroom where they were expected to observe for at least 20 hours over the course of the semester. Over a six- to eight-week period, students visit their assigned classroom one day a week for approximately 2-4 hours each visit. For the two live sections of the course where the class meets face-to-face, the traditional field experience includes a weekly one-hour meeting dedicated to discussion/activities related to the field experience.

Included in this study was also a web-based section of the course that completed a traditional field experience (where they observed for 20 hours at a local school). Since there was no live, face-to-face weekly meetings dedicated to the field experience (as there was for the face-to-face courses), all discussion/activities related to the field experience took place via online discussions. The instructor often provided students with focus questions/topics for the field observations and required students to post and discuss responses to those questions in an online discussion forum.

Virtual Field Experience. Students completing the virtual field experience were enrolled in a web-based section of Introduction to Educational Psychology and were using the video-based classrooms in the Inquiry Learning Forum (ILF) as the foundation of their field experience. The ILF (http://ilf.crlt.indiana.edu) has many components including a library of lesson plans and web resources, a collection of inquiry-based professional development "labs," a collaboratory with a variety of public and private asynchronous discussion forums, and a collection of video-based classrooms. Throughout the semester, students were required to view several video cases from the ILF collection and completed a variety of activities related to the cases they view and the concepts they are learning in the course

Instrument

The questionnaire for students completing the traditional field experience contains 37 questions, while 38 questions were offered to those completing the virtual field experience. The majority of questions were close-ended, asking students to select an appropriate response from given choices. Questions related to the opportunities of the virtual field experience over the traditional field experience (Research Question #1) addressed the following topics: reflection time, rewatching of video clips, discussion with classmates, observation of various student and teacher characteristics, and links to class content. A series of questions related to the authenticity of the field experience (Research Question #2) asked about the field experience's effect on students' understanding of teaching, their decision to teach, their view of the teacher as a mentor, their discussions with the classroom teacher, and their involvement in the classroom. Student's satisfaction with the field experience (Research Question #3) was addressed by questions asking about what they learned from the teacher, whether it added value to their course, if they enjoyed the experience, and whether they got ideas for their future classroom. The final questions on each questionnaire were open-ended, allowing students to comment about the perceived strengths and limitations of their field experience format (Research Question #3).

Design and Data Analysis

This study was a quasi-experimental study with one independent variable with two levels – type of field experience (traditional and virtual). The dependent variables were students' perceptions of and experiences in the field experience, which were operationalized by two sets of items on the questionnaire. The objective items on the questionnaire were analyzed with a MANOVA and follow-up ANOVAs. The two open-ended questionnaire items were analyzed to identify emerging trends.

Procedure

A list of student email addresses was obtained from course instructors. Students received an email message inviting them to participate in the study and providing them with a link to the online questionnaire. Two reminder emails were sent to all invited participants during the following week. Although it was not the intent of the researchers, the instructors of the two live sections engaged in a discussion with their students about the virtual field experience and the purpose of this research study prior to their participation.

Results

Demographic data

Forty-two of the 86 students in the traditional group and eight of the 20 students in the virtual group completed the online questionnaire. Response rates were 49% and 40%, respectively. Participants who completed the traditional field experience included 28 women and 14 men, and their mean age was 21.14 years. Participants in the virtual field experience group included 3 women and 5 men, and their mean age was 21.38 years. Most students who completed the questionnaires were sophomores: 83.3% in the traditional group and 62.5% in the virtual group.

At the time of questionnaire administration, students in the traditional group reported that they had completed an average of 15.27 hours of observation in the classroom and expected to spend a total of 21.26 hours by the end of the semester. Students completing the virtual field experience reported spending 12.69 hours watching the video cases at the time of questionnaire administration and expected to spend a total of 15.81 hours by the end of the semester.

Comment: Should we add something here that mentions how much time they spent observing at the point of the survey administration and how much they expected to do during the semester? I was going to add that information, but there was lots of missing data on this one from the VFE students…do you know why that is? Can you check the original data?

Questionnaire Items

To examine students' perceptions of and experiences in each field experience format, a MANOVA was conducted on relevant questions from the questionnaire. The MANOVA indicated that the type of field experience a student completed (virtual or traditional) had a significant effect on student's perceptions of and experiences in the field experience, $F(26, 21) = 3.12$, $p = .005$, $0^2 = .21$. Follow-up ANOVA revealed four significant items which will be discussed below in relation to each specific research question.

Research Question #1: Do students completing the virtual field experience recognize and take advantage of opportunities that the traditional field experience may not offer? All students in the virtual field experience who completed the questionnaire reported that they did re-watch video clips or portions of video clips when completing the video-based assignments (6 students reported re-watching clips occasionally, 2 reported doing so frequently).

Students were also asked to indicate on a scale how much time they spent reflecting on each hour of observation (including completing assigned activities and discussing observations with friends). On the scale, 1 = no time; 2 = less than .5 hours; 3 = .5 to 1 hour; 4 = 1 to 1.5 hours; 5 = 1.5 to 2 hours; 6 = 2 to 2.5 hours; 7 = 2.5 – 3 hours. Students completing the virtual field experience reported spending more time reflecting on their observations (M = 4.50, SD = 1.85) than students completing the traditional field experience (M = 3.26, SD = 1.25), $F(1, 45) = 5.87$, $p = .019$, $0^2 = .11$. Students completing the virtual field experience on average reported spending 1 – 2 hours reflecting on each hour of observation, whereas those completing the traditional field experience generally spent less than 1 hour reflecting on each hour of classroom observation. It is also interesting to note that students completing the traditional field experience expected to spend more time observing in the classroom (M = 21.26, SD = 3.06) than the virtual field experience students expected to spend watching the video cases (M = 15.81, SD = 5.25), $F(1, 45) = 21.14$, $p < .001$, $0^2 = .32$.

Other questionnaire items related to this research question addressed topics including discussion with classmates, observation of various student and teacher characteristics, and relating observations to class content. Student responses are on a 5-point Likert-type ranging from Strongly disagree (1) to Strongly agree (5). No other items related to this research question indicated a significant difference; these items are included in Table 2.

Table 2 *Means and standard deviations for questionnaire items related to research question #1*

	Virtual		*Traditional*	
	M	*SD*	*M*	*SD*
Throughout the field experience, I was able to engage in meaningful discussion about my observation with my classmate(s).	4.12	.64	3.80	.91
The field experience gave me the opportunity to observe various developmental levels of students.	3.88	.64	3.98	.95
The field experience gave me the opportunity to observe different aspects of social development in students.	3.50	.53	3.92	.73
The field experience gave me the opportunity to observe issues related to inclusion (as it relates to special education).	2.38	.92	3.15	1.29
The field experience gave me the opportunity to observe different teaching styles.	4.38	.52	3.60	1.10
The field experience gave me the opportunity to observe different classroom management techniques.	3.88	.64	3.82	1.01
The field experience gave me the opportunity to observe different approaches to assessment and evaluation.	3.88	.64	3.62	.84
The field experience gave me the opportunity to observe various teaching and learning theories “in action” (e.g., social learning theory, information processing theory, constructivism).	4.12	.99	3.65	1.03
I often observed situations/events in the classroom(s) that I could relate to specific concepts I was learning in class.	4.25	.71	4.00	.82
I often reference observations from my field experience in discussions related to course content.	4.12	.83	3.80	.85
I feel that the field experience helped me learn the course content.	4.12	.99	3.62	.84

Note. Means reflect student responses on a five-point Likert-type scale ranging from Strongly disagree (1) to Strongly agree (5).

Research Question #2: *Do students completing the virtual field experience view it as authentic? Do they differ from those completing the traditional field experience in their ability to “get a feel” for being in the classroom?* Students participating in the traditional field experience reported engaging in more meaningful discussions with the classroom teacher(s) (M = 3.08, SD = .89) than did those completing the virtual field experience (M = 1.88, SD = 1.24), $F(1, 46) = 10.60$, $p = .002$, $0^2 = .19$. The scale for this item asked students to rate the frequency of discussions as (1) “Never,” (2) “A couple of times,” (3) “Occasionally,” or (4)

"Frequently." While the majority of the virtual field experience respondents (62%) reported that they never engaged in discussion with the classroom teacher, only one traditional field experience student (2%) reported not having engaged in any such discussions.

Other questionnaire items related to this research question did not indicate significant differences; these items are included in Table 3.

Table 3 *Means and standard deviations for questionnaire items related to research question #2*

	Virtual		*Traditional*	
	M	*SD*	*M*	*SD*
The field experience helped me to better understand what teaching is really like.	3.62	1.06	4.00	1.06
The field experience impacted my decision to be a teacher (either reinforced it or led you to change your mind).	3.38	.74	3.78	1.10
I view the teacher(s) I am observing as a mentor.	3.38	.92	3.60	1.17
I felt involved in the classroom(s) I was observing.	2.75	.71	3.70	1.30

Note. Means reflect student responses on a five-point Likert-type scale ranging from Strongly disagree (1) to Strongly agree (5).

Research Question #3: Do students completing the virtual and traditional field experiences differ in terms of their satisfaction with and value gained from their field experience? What strengths and weaknesses do students perceive for their specific field experience format? Students completing the traditional field experience rated their time spent observing classroom(s) as significantly more enjoyable (M = 4.20, SD = 1.11) than those completing the virtual field experience (M = 3.12, SD = 1.13) $F(1, 46) = 6.19$, $p = .017$, $0^2 = .12$.

Interestingly, 75% of the respondents completing the traditional field experience disagreed (over 50% strongly disagreed) that they would have preferred to complete the field experience in the virtual format. Similar patterns emerged when asked if they felt they would have learned more in a virtual field experience.

Other questionnaire items related to this research question did not indicate significant differences; the items are included in Table 4.

Table 4 *Means and standard deviations for questionnaire items related to research question #3*

	Virtual		*Traditional*	
	M	*SD*	*M*	*SD*
I feel that I learned a great deal from the teacher(s) I observed.	3.75	.71	3.85	1.14
The field experience added value to the Introduction to Educational Psychology course.	4.00	.53	3.88	.89
I observed teaching strategies that I would like to use in my future classroom.	4.50	.53	3.78	1.05
I observed teaching strategies that I would not use in my future classroom.	4.12	1.36	3.55	1.01
Through my field experience, I gained ideas that I will be able to implement in my future classroom.	4.38	.52	3.98	1.17
The teacher(s) I observed modeled effective teaching practices.	3.75	.71	3.92	1.07

Note. Means reflect student responses on a five-point Likert-type scale ranging from Strongly disagree (1) to Strongly agree (5).

Open-ended questions

Students were also asked to respond to open-ended questions regarding what they perceived to be the strengths and limitations of their field experience format. The trends that emerged for each field experience format are identified in Table 5 and discussed below with representative participant responses included.

Table 5 *Summary of strengths and limitations as identified by respondents.*

	Strengths	Limitations
Traditional	• one-on-one interaction with students and teachers • getting a feel for what a classroom is really like • hands-on experience • observing teaching and management strategies • influencing teaching as a future career • exposure to student diversity	• meeting 20-hour requirement • distance and transportation • problems related to the structure of field experience • observing only one classroom and one teacher • limited role in experience • placed in classroom unrelated to subject area • no diversity related to students and schools • placed with poor classroom teacher
Virtual	• flexibility in scheduling • no need to drive • quality of video clips	• lack of classroom feel • lack of interaction with teacher and students • limited scope of observation

Traditional field experience group. Of the 42 respondents who participated in the traditional field experience, 35 of them completed the open-ended item asking about the strengths of the traditional field experience. Based on their responses, a number of perceived strengths were identified: (a) one-on-one interaction with students and teachers, (b) getting a feel for what a classroom is really like, (c) hands-on experience, (d) observing teaching and management strategies, (e) influencing teaching as a future career, and (f) exposure to student diversity. Representative responses regarding the strengths of field experience with the number of similar responses are shown below:

- One-on-one interaction with students and teachers ($f = 21$): "I love that fact that I get to personally interact with a real child." "It is nice to actually be there and be able to interact with the teacher."
- Getting a feel for what a classroom is really like ($f = 18$): "You get a real feel for what could be your classroom in a couple years." "It is like the old saying a picture is like a thousand words, well an actual classroom is the same way. A future teacher cannot get the whole picture without actually being there."
- Hands-on experience ($f = 7$): "I feel that the strengths of observing in local schools are that I was able to get a first hand experience of the classroom atmosphere through observing."
- Observing teaching and management strategies ($f = 5$): "…learning how classroom management works…" "Seeing everything, good and bad and how it is dealt with."
- Influencing teaching as a future career ($f = 4$): "It reinforced my ambitions to be a teacher." "This experience has given me a positive outlook on teaching."
- Exposure to student diversity ($f = 3$): "It opened my eyes to different socio-economic backgrounds" "…it gives us the opportunity to realize what children's learning styles are at that age…"

Thirty-one of the 42 respondents in the traditional group responded to the question asking about perceived limitations of the traditional field experience. Some of the limitations identified include: (a) meeting the 20-hour requirement, (b) distance and transportation, (c) problems related to the structure of field experience, (d) observing only one classroom and one teacher, (e) limited role in experience, (f) placed in classroom unrelated to subject area, (g) no diversity related to students and schools, and (h) placed with poor classroom teacher. Interestingly, 8 students (19% of respondents) reported that there was no limitation in the field experience. Below are examples of student responses related to the limitations of the traditional field experience with the number of similar responses indicated:

- Meeting the 20 hour requirement ($f = 7$): "I somewhat feel that twenty hours out in the field is pointless." "Make a minimum number of visits with a shorter time span (say 6 visits for 2 hours)."
- Distance and transportation ($f = 6$): "A limitation for me personally, is trying to have an effective form of transportation that runs smoothly with the time schedules."
- Problems related to the structure of field experience ($f = 4$): "When you only meet once a week for a short period you only are able to observe a limited amount."

- Observing only one classroom and one teacher (*f* = 4): "Sitting in the same classroom with the same teacher giving the same lesson plan three times in a row is not helpful."
- Limited role in experience (*f* = 4): "I felt more like hired help then someone there that wanted to learn more about teaching."
- Placed in classroom unrelated to subject area (*f* = 3): "The main limitations I saw are the limitations of finding a placement in a specific area of the student's study. I was placed in a Physics class. I will not be teaching Physics."
- No diversity related to students and schools (*f* = 3): "On the other hand I don't get to observe classrooms all over the US with different cultures."
- Placed with poor classroom teacher (*f* = 2): "I think that the only limitation would be a teacher that was not using this opportunity to best benefit the student who is observing."

Virtual field experience group. Of the eight respondents who completed the virtual field experience, all of them completed the open-ended item asking about the strengths of the virtual field experience. Strengths of the virtual field experience identified by students include: (a) flexibility in scheduling, (b) no need to drive, and (c) quality of video clips. Below are representative responses for each strength with the frequency of similar responses indicated:

- Flexibility in scheduling (*f* = 4): "They allow you to complete assignments and observation on your own schedule."
- No need to drive (*f* = 2): "Students do not have to drive to a surrounding school."
- Quality of video clips (*f* = 2): "Another strength is that we are provided with the teacher reflections, flexible observation clips, and student works. These items are rarely, if ever, available in a classroom observation."

All eight students in the virtual group completed the question about perceived limitations of the virtual field experience. Among the limitations identified were: (a) lack of classroom feel, (b) lack of interaction with teacher and students, and (c) limited scope of observation. Representative responses and frequencies of similar responses are as follows:

- Lack of classroom feel (*f* = 3): "You don't get to sit in on an actual classroom. In fact I think that if there is one class that shouldn't be web based it is this one."
- Lack of interaction with teacher and students (*f* = 2): "To actually see a child and how they react to a teacher is much better than just 'trying' to see him/her on a computer screen."
- Limited scope of observation (*f* = 2): "You don't get to see all aspects all the time. You really only get one viewpoint and its portrayed as the teacher wants you to see it."

Discussion

It is interesting to note that some of the most prevalent strengths mentioned by those completing the traditional field experience are the most commonly mentioned weaknesses of the virtual field experience and vice versa. The three most frequently listed strengths of the traditional field experience are the opportunity for one-on-one interactions with students and the teacher, the ability to "get a feel" for the classroom and what teaching is really like, and having the opportunity to gain hands-on experience with teaching. Students completing the traditional field experience emphasized the importance of being physically present in a classroom and being part of that environment. Virtual field experience students also recognize the value of being in the classroom which is demonstrated by what they identified as the two most prevalent limitations of the virtual field experience –inability to "get a feel" for the classroom and lack of interactions with the students and teacher.

Similarly, the top strengths of the virtual field experience are the most commonly mentioned weaknesses of the traditional field experience. The virtual field experience students appreciated the fact that they were not tied to a schedule allowing them to complete observations whenever it was convenient, and that they did not have to travel to get to their observations. Similar topics were addressed by the traditional field experience students who reported having difficulty meeting the 20-hour observation requirement and traveling to their field placement. Especially as students are being placed at increasingly distant locations, the logistics of getting to that location to complete 20 hours of observation is becoming problematic.

One of the commonly found limitations of traditional field experiences is that students often do not spend adequate time processing and reflecting on their observations, thereby not gaining as much as they perhaps could (or should) from the experience (Feiman-Nemser & Buchmann, 1985; Goodman, 1986; Johnston, 1994). The findings of this study suggest that the virtual field experience may address this concern. The virtual field experience students reported spending at least as much time reflecting on the video-based cases as they did watching them, which is contrasted with the traditional field experience students who spent less time reflecting than they did observing. It appears that students in the two field experience formats demonstrate different patterns of activity. While the virtual field experience students spend more time reflecting and less time watching the video cases, the traditional field experience students spend more time observing in the classroom and less time reflecting on those observations. It is important for preservice teachers to understand the value of reflection and to be encouraged to become reflective practitioners; a successful field experience needs to promote such practice (e.g., Freiberg, 1995; Norlander-Case, Reagan, Case, 1999; Posner, 1996).

It is also interesting to note that all of the respondents participating in the virtual field experience reported re-watching portions of the video cases, which is obviously not possible when observing a live classroom. It is possible that being able to review video clips at various stages in the thought process promotes more and deeper reflection. Future research needs to investigate the quality of reflection done by virtual field experience students and work to identify the qualities of the virtual field experience that promote increased reflection.

Another key finding from this study is that students who participated in the virtual field experience did not enjoy their experience as much as those who completed the traditional field experience. It is not entirely surprising that students may not find sitting in front of a computer screen as enjoyable as interacting with students and the teacher in an actual classroom. However, this issue is one that should be acknowledged by teacher educators and taken into consideration when discussing field experience formats. Future research should investigate what qualities of the virtual field experience may impact student enjoyment and satisfaction level, so efforts can be made to address those issues and make the experience more enjoyable.

Another limitation of the virtual field experience is the lack of opportunity to engage in meaningful discussions with the classroom teacher. While all but one of the traditional field experience students reportedly engaged in at least one meaningful discussion with the classroom teacher, the majority of virtual field experience respondents did not engage in any such discussions. This finding is not surprising given the minimal opportunity for such interactions in the virtual environment. While there are discussion forums related to each video classroom open to all Inquiry Learning Forum members (including the classroom teachers in the videos), these discussions are not very active and it is possible (or likely) that the participating teacher may not see and/or respond to posted messages in a timely fashion (if at all). Providing extensive notes about each teacher's reflections on a given lesson may compensate for some of this lack of interaction, but other avenues for communication with the teacher should also be investigated.

Many of the virtual field experience students also commented that they did not get a sense of what it was like to be a teacher in a classroom. As stated previously, there can be a variety of objectives for any given field experience, but if "getting a feel" for the classroom and interacting with students and teachers is the primary goal of the field experience, the virtual field experience may be limiting. Various multimedia components can be used to create a rich learning environment, however nothing can replace being in an actual classroom to see, hear, and smell what it is like to be a teacher. One of the traditional field experience students said it best, "It is like the old saying a picture is like a thousand words, well an actual classroom is the same way. A future teacher cannot get the whole picture without actually being there."

Although the findings of this study have implications for those considering alternatives to the traditional field experience, there are several limitations of this study. The biggest limitation of this study is the small number of participants (n=8) participating in the virtual field experience. Another limitation of the current study was that there was no random assignment. Since students were already enrolled in the courses, the researchers were not able to randomly assign them to the groups. Finally, the questionnaire was administered to students prior to their completing the entire field experience. It was administered about two-thirds of the way through the semester; and although students were well into their field experiences at that point, it would have been preferable to administer the questionnaire at the conclusion of the semester.

Conclusion

As was stated previously, the purpose of this study was not to identify which field experience was "better," but rather to better understand the students' experiences in and perceptions of virtual and traditional field experiences. It is hoped that the information provided here will be of use to individuals who are considering a virtual field experience as an alternative to the traditional field experience. It has become clear that the overall goals and objectives for a specific field experience must be the focus when field experience options are being explored. Additionally, when comparing various formats, it is important to consider both the experience itself and the learning that occurs as a result. The current study addressed only the student experience, and further research is needed to investigate the learning related to each type of field experience. Each format has its strengths and limitations, and perhaps it is possible to create a hybrid experience that builds on the strengths and minimizes the weaknesses of each format.

References

Aiken, I. P., & Day, B. D. (1999). Early field experiences in preservice teacher education: Research and student perspectives. *Action in Teacher Education, 21*(3) 7-12.

Atkins, S. L. (1998). Windows of opportunity: Preservice teachers' perceptions of technology-based alternatives to field experiences. *Journal of Computers in Mathematics and Science Teaching, 17*(1), 95-105.

Barab, S. A., MaKinster, J. G., Moore, J. A., & Cunningham, D. J. (2001). Designing and building an on-line community: The struggle to support sociability in the Inquiry Learning Forum. *Educational Technology Research & Development, 49*(4), 71-96.

Bonk, C. J., Malikowski, S., Angeli, C., & East, J. (1998). Web-based conferencing for preservice teacher education: Electronic discourse from the field. *Journal of Educational Computing Research, 19*(3), 269-306.

Brooks, D. & Kopp, T. W. (1990). Technology and teacher education, in: W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 498-513). New York: Macmillan.

Chaney-Cullen, T., & Duffy, T. M. (1999). Strategic teaching framework: Multimedia to support teacher change. *Journal of the Learning Sciences, 8*(1), 1-40.

Cruickshank, D. R., & Armaline, W. D. (1986). Field experiences in teacher education: Considerations and recommendations. *Journal of Teacher Education, 37*(3) 34-40.

Feiman-Nemser, S., & Buchmann, M. (1985). Pitfalls of experience in teacher preparation. *Teachers College Record, 87*(1), 53-65.

Frieberg, H. J. (1995). Promoting reflective practices. In G. A. Slick (Ed.), *Emerging trends in teacher preparation: The future of field experiences* (pp. 25-42). Thousand Oaks, CA: Corwin Press, Inc.

Garrett, J. L., & Dudt, K. (1998). *Using video conferencing to supervise student teachers*. In Proceedings of SITE 98: Society for Information Technology & Teacher Education International Conference (9th, Washington, D.C.).

Goodman, J. (1986). Making early field experience meaningful: a critical approach. *Journal of Education for Teaching, 12*(2), 109-125.

Gruenhagen, K., McCracken, T., & True, J. (1999). Using distance education technologies for the supervision of student teachers in remote rural schools. *Rural Special Education Quarterly, 18*(3/4), 58-65.

Hannah, C. L. (1995). Self-study: Students evaluate the use of video "cases" in an educational psychology course with a field component. *Journal of Technology and Teacher Education, 3*(2/3), 267-279.

Hannah, C. L., & Abate, R. J. (1992-1993). Survey on faculty use of interactive videodisc technology in teacher education. *Journal of Educational Technology Systems, 21*, 321-332.

Hopkins, S. (1995). Using the past; guiding the future. In G. A. Slick (Ed.), *Emerging trends in teacher preparation: The future of field experiences* (pp. 1-9). Thousand Oaks, CA: Corwin Press, Inc.

Johnston, S. (1994). Experience is the best teacher; or is it? An analysis of the role of experience in learning to teach. Journal of Teacher Education, 45(3), 199-208.

Lambdin, D. V., Duffy, T. M., & Moore, J. A. (1996). A hypermedia system to aid in preservice teacher education: Instructional design and evaluation. In Proceedings of selected research and development presentations at the 1996 national convention of the Association for Educational Communications and Technology (18th, Indianapolis, IN).

Li, X., & Zhang, M. (2000). *Effects of early field experiences on preservice teachers' efficacy beliefs – A pilot study*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Mason, C. L. (2000). Online Teacher Education: An Analysis of Student Teachers' Use of Computer-Mediated Communication. *International Journal of Social Education, 15*(1), 19-38.

McDevitt, M. A. (1996). A virtual view: Classroom observations at a distance. *Journal of Teacher Education, 47*(3), 191-195.

Mclntyre, D. J., & Pape, S. L. (1993). Interactive video for teacher education. *The Teacher Educator, 28*(3), 2-10.

Metcalf, K. K., & Kahlich, P. A. (1996). Laboratory experiences as transition from campus to field. In D. J. McIntyre, & D. M. Byrd (Eds.), *Preparing tomorrow's teachers: The field experience* (pp. 97-114). Thousand Oaks, CA: Corwin Press, Inc.

Moffett, D. (2001). *Using the Internet to enhance student teaching and field experiences*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL.

Norlander-Case, K. A., Reagan, T. G., & Case, C. W. *The professional teacher: the preparation and nurturance of the reflective practitioner.* San Francisco, CA: Jossey-Bass.

Pape, S. L., & McIntyre, D. J. (1993). *Student reactions to using video cases for evaluation of early field experiences*. Paper presented at the summer workshop of the Association for Teacher Educators, Pittsburgh, PA.

Paese, P. C. (1985). Assessment of a teacher education program based on student intern performance. *Journal of Teaching in Physical Education, 5,* 52-58.

Paese, P. C. (1996). Contexts: Overview and framework. In D. J. McIntyre, & D. M. Byrd (Eds.), *Preparing tomorrow's teachers: The field experience* (pp. 1-7). Thousand Oaks, CA: Corwin Press, Inc.

Posner, G. J. (1996). *Field experience: a guide to reflective teaching*. White Plains, NY: Longman Publishers USA.

Representations and Causal Models:
Instructional effects of text, narration, virtual manipulation, and sound effect in a multimedia tutorial

Bob Hoffman
San Diego State University

Ray Adair
School of Creative and Performing Arts

Andrea Cascia
San Diego City Schools

Abstract

Subjects viewed a computer-based instructional multimedia presentation showing how to raise a canal barge through a lock. Along with animation, subjects received verbal instructions for the steps involved in the process as either narration or text. Some subjects were required to use the computer mouse to virtually manipulate screen objects such as images of the barge and lock valves and gates, while others viewed the animated steps without manipulating the objects, using only a "Next" button to continue. Some subjects heard realistic sound effects associated with movements of screen objects, while others did not. We found no significant differences between the scores of those who heard narrated verbal instructions and those who read the same instructions as text, on a retention test, in matching labels to images of objects, or in correctly identifying solutions to transfer problems. We also found no significant differences on retention, matching, or transfer tests between subjects who manipulated screen objects and those who did not, though there was a significant interaction between verbal presentation modality (narration or text) and virtual manipulation. Subjects who heard realistic sound effects associated with screen actions performed better on the matching test than subjects who heard none, but we found no significant differences in performance between the two groups on the retention and transfer tests. These results help instructional designers make more informed decisions about the value added to learning outcomes relative to the development costs of multimedia features such as virtual manipulation and sound effects.

Introduction

Educational project designers increasingly employ multimedia representations in educational products ranging from on-line training and education to electronic reference works. Multimedia representations include combinations of text, still images, animation and motion pictures, and sound effects. Sometimes designers provide interactive control over various aspects of these representations, letting learners directly or indirectly manipulate screen elements, including visual images and diagrams, and providing visual and/or aural feedback.

Developers must weigh the educational value added by a particular representation against the relatively high production costs for imagery and sound, compared with those for text. For example, using even a simple animation where a still image or text description might serve learners almost as well might not be an efficient use of resources. At the very least, designers want to understand the degree to which various representations or combinations of representations are likely to facilitate or interfere with learning, in various contexts.

This experiment employed text and computer-generated representational graphics and sound to examine memory and transfer effects in graduate students learning about processes and causal relationships involved in moving a barge from one level to another through a canal lock. The independent variables consisted of eight combinations of various representational conditions common in multimedia instruction. These conditions were: 1) verbal presentation (aural narration or text); 2) transformation of key visual objects (learner controlled or computer controlled); and 3) auditory feedback associated with canal operations (sound effects or no sound effects).

The study examined the effects of these combinations on three dependent variables, retention, matching, and transfer. A common goal for most tutorial instruction is to help learners retain conceptual information, such as the sequence of events or operations in a process. Tutorials also may help learners remember facts, such as matching images of objects with the names of those objects. Far transfer involves applying learned principles to analyze or solve novel problems. Grasping the principles in a worked problem is a necessary prerequisite to applying those principles in new circumstances.

The experiment tested three hypotheses, based on assumptions gleaned from the literature on cognitive psychology. First, we predicted that subjects who completed a lesson in which explanatory text accompanied visual transformations of canal objects, such as the movements of barges and gates, would learn as much or more than those who heard the same explanation as narration—a spoken soundtrack. While Mayer and Moreno (1998) found that subjects viewing fixed-paced animated sequences learned more from accompanying narration than from text, we hypothesized that text might prove more useful when learners controlled the pacing of animated sequences, allowing them to better relate the text to the animation. Second, we predicted that subjects who exercised control of representational objects— such as the canal gate and barge—through mouse movements would understand and remember more about canal operations than subjects who "passively" viewed similar transformations of these representational objects in pre-determined sequences controlled by the computer. We based this hypothesis on assumptions that object manipulation results in increased attentiveness, in turn resulting in learner's encoding more relevant information about lock operations, thus facilitating deeper mental processing and more thorough integration of new and existing knowledge. The hypothesis was also based on the idea of the intimate relationship between action and perception. For example, drivers and passengers remember different information from the same virtual journey (Brooks, et al, 1999).

Third, we hypothesized that learners who heard naturalistic sound effects associated with animation depicting canal operations would understand and remember more than learners who heard no sound effects. We based this prediction on assumptions that sound effects would cue learners to attend to contemporaneous visual and verbal information (Alwitt, Anderson, & Lorch, 1980), and that the sound effects themselves provide a redundant source of content information. For example, a slow, deep creak might help learners distinguish a heavy, wooden gate from the higher-pitched squeak of a smaller valve.

Method

Participants and design

Three hundred forty-five graduate students participated as subjects in this study. They were enrolled in a Masters Degree program in educational technology or in a teacher credential program at San Diego State University. Identical pre- and post-tests measured retention, matching, and transfer. We used performance on the pre-test to screen out subjects who already understood how canal locks worked, and therefore presumably would not benefit from instruction, eliminating those who responded correctly to more than 2 out of 5 retention/matching items, and more than 4 out of 8 transfer items. We eliminated 215 subjects in this way, leaving 130 to be randomly assigned by a computer algorithm to one of eight treatments (Table 1) at the time they began the experiment.

Materials and apparatus

The study employed an animated tutorial of a barge moving from a lower to a higher level in a canal by way of a lock (see Figure 1). Informal studies conducted during the design and development of this tutorial indicated that, while many people have a general idea that canal locks move boats up and down, most couldn't describe the mechanics of these movements, or the various components involved, such as gates, swing beams, valves, water levels, and the barge itself.

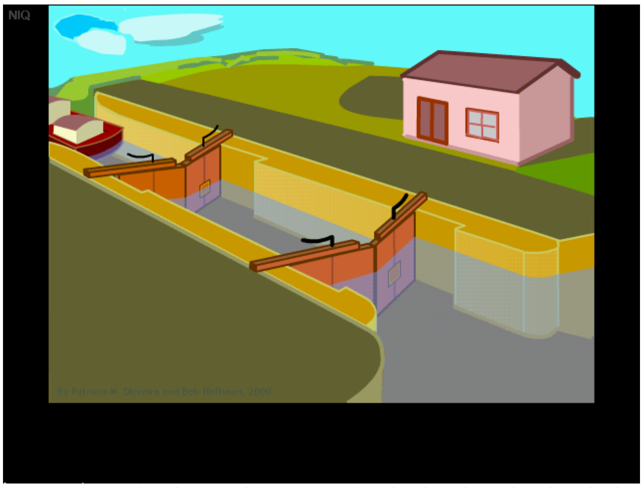


Figure 1. Opening frame of the instructional animation, showing the barge (upper left), the canal, and the lock with gates, swing beams, sluice valves and sluice valve keys. Subjects move through successive steps by using their mouse to manipulate these screen objects as directed.

We created eight different treatments of the animation, based on a 2x2x2 factorial design (Table 1): verbal instruction as narration or text; virtual manipulation versus no manipulation; and sound effects versus no sound effects.

Table 1 *Treatments for the canal lock experiment.*

Treatment	Verbal Instruction	Manipulation	Sound Effects
1	Narration	Virtual	Sound Effects
2	Text	Virtual	Sound Effects
3	Narration	None	Sound Effects
4	Text	None	Sound Effects
5	Narration	Virtual	None
6	Text	Virtual	None
7	Narration	None	None
8	Text	None	None

Verbal instruction (the words that described the steps for moving a barge through a canal lock) was presented either as text or as narration. Text set in white letters on the black Web page background (Figure 2), was positioned just beneath the animation to minimize any possible split-attention effect caused by spatial separation (Chandler & Sweller, 1992; Mayer & Moreno, 1998). We designed a single set of verbal instructions that would work for all eight treatments. The instructions divided the process into five discreet "steps" for moving the barge from the lower to the upper level of the canal (Figure 2). Subjects in the text condition could take as much time as they wished to re-read the text before going on to the next frame, but subjects in the narration condition could not replay the narration.

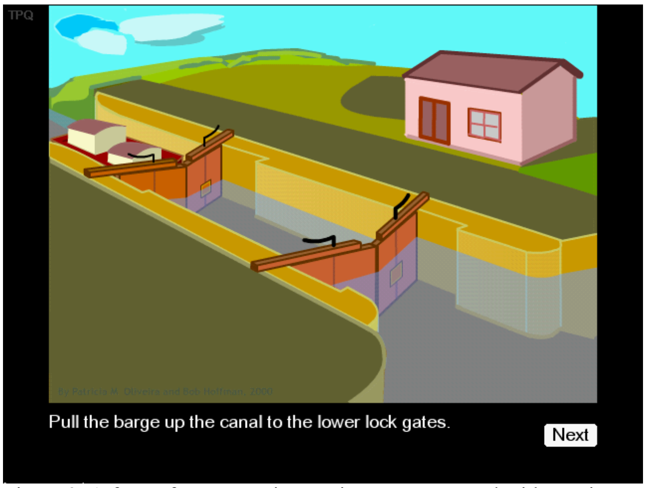


Figure 2. A frame from a non-interactive treatment, and with text instead of narration. Subjects continue through the steps by clicking on the "Next" button in the lower right corner.

Subjects assigned to the virtual manipulation condition (Figure 1) were obliged to use their mouse to click and drag graphic elements when prompted by the verbal instructions and a hand-shaped cursor. For example, they were instructed to pull the barge through the lock and to open and close gates and valves. Subjects in the non-manipulation condition (Figure 2) watched the barge move and the gates and valves open and close, but needed only to click a "Next" button to view the next step. Subjects in neither condition could go back to review a previous step. As soon as they completed the required action or clicked on “Next,” the program moved automatically to presentation of the next step.

Half of the treatments featured realistic sound effects of squeaking valve handles, creaking gates, and water rushing through valves, accompanying the corresponding animation, while the other half presented no synchronized sound effects. All treatments featured background sounds of water lapping against the sides of the canal, in part so that subjects who did not receive the instructionally-related sound effects or narrations would not be distracted by wondering whether they should be hearing something, since all subjects wore earphones and were asked to adjust the sound level to suit their preference at the beginning of the experiment.

All treatments featured exactly the same animation, including a blinking “highlight” effect to draw attention to relevant objects as they were mentioned in the accompanying narration or text.

The experiment started with a title screen, a brief description of the experiment, a sound test and adjustment screen, and a demographics page which asked subjects to report their age, gender, grade in school, and whether or not they used a computer at home. We reasoned that subjects with more computer experience might be more comfortable with the interactive multimedia environment, experience less cognitive load, learn more, and perform better on the test items.

We then presented a short pre-test, on-screen, consisting of five items designed to determine an initial level of content knowledge and to provide a benchmark for post-testing retention, ability to match labels with objects, and transfer. Each question was presented on a separate screen.

After answering or passing on the pre-test questions, subjects came to a screen with instructions for beginning the treatment. The program then switched each subject to a randomly assigned treatment. Following the treatment, the same items (with some minor changes in the wording of instructions or stems) constituted the post-test, with the difference that the program required an attempt at a response on each item before allowing a subject to continue.

Following the last post-test question, the program sent the subject’s data (including elapsed times for and answers to each pre- and post-test question, and the identity and duration of the treatment) to an on line database, and sent the subject’s browser to a Web page inviting the subject to view their own results on the post-test along with the collective results of all subjects up to that moment. The results page identified the treatment the subject had received, described all eight treatments, displayed the user’s score on the post-test along with the average post-test scores and number of subjects for each treatment up to that moment, and presented a short discussion of the experiment and what the results might mean.

Procedure

Experimenters prepared subjects’ usual computer classrooms in advance or, in a few cases, invited their class to a prepared computer classroom. Each subject sat at a 15" (13.8" viewable) display set to 640x 480 pixels per inch resolution. Each subject wore adjustable-volume headphones and controlled the program with a conventional, mechanical one-button computer mouse and mouse pad. When subjects were comfortably seated at their workstations, experimenters asked them to begin by following the instructions on their computer screen (see above).

Results

We conducted an ANOVA with verbal presentation (narration vs. text), manipulation (virtual manipulation vs. no manipulation), and sound effects (vs. no sound effects) as the three independent variables and gains in scores between pre-tests and posttests on retention, matching, and transfer items as the dependent variables.

Verbal Presentation On Retention, Matching, And Transfer

We predicted that subjects who read text accompanying animation in self-paced instruction would recall the sequence of steps and label the parts of the lock as well as or better than subjects who listened to narration accompanying animation. We also expected subjects in the text condition to be able to transfer their understanding to make predictions based on underlying principles as well as or better then subjects in the narration condition. The ANOVA failed to indicate a significant effect of verbal presentation on retention, matching, or transfer. On retention, $M = 3.47$ ($SD = 1.96$) for the narration group and $M = 3.62$ ($SD = 1.88$) for the text group; on matching, $M = 3.11$ ($SD = 1.68$) for the narration group and $M = 3.53$ ($SD = 1.54$) for the text group; on transfer, $M = 4.91$ ($SD = 2.61$) for the narration group and $M = 4.78$ ($SD = 3.11$) for the text group.

Virtual Manipulation On Retention, Matching, And Transfer

The virtual manipulation group used the computer mouse to open and close gates and valves and pull barges into position in order to move to the next step. We predicted they would better retain the sequence of steps, match pictured objects with labels, and transfer the causal model presented to novel situations, than would subjects who merely observed the animated actions and clicked a “Next” button to continue.

The non-manipulation group achieved slightly higher mean gains than the virtual manipulation group in retention of the sequence of steps, labeling objects, and the four transfer items, but this effect was not significant on any of the measures. On retention, M = 3.49 (SD = 2.05) for the virtual manipulation group and M = 3.57 (SD = 1.79) for the no manipulation group; on matching, M = 3.16 (SD = 1.63) for the virtual manipulation group and M = 3.41 (SD = 1.63) for the no manipulation group; on transfer, M = 4.84 (SD = 2.76) for the virtual manipulation group and M = 4.87 (SD = 2.90) for the no manipulation.

Presentation x Virtual Manipulation On Matching.

There was a significant interaction (Figure 3) between presentation (narration vs. text) and virtual manipulation (vs. no manipulation). Subjects who read text explanations performed better in the no manipulation condition and subjects who listened to narration did better when they manipulated the animation, F(1, 122) = 4.338, MSE = 2.458, p = .039.

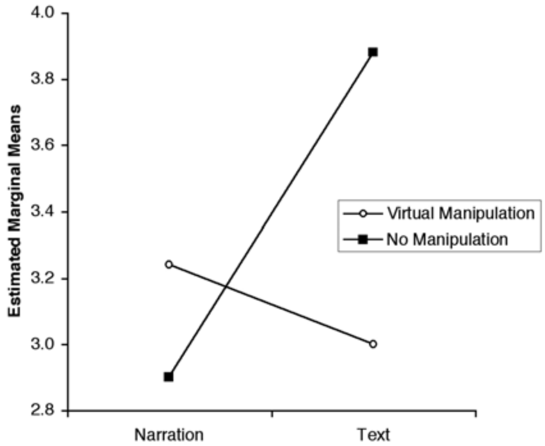


Figure 3. Estimated marginal means for interaction effect between presentation (narration and text) and virtual manipulation (vs. no manipulation) groups on matching labels with pictured objects. M = 2.90 (SD = 1.95) for the narration with virtual manipulation group; M = 3.00 (SD = 1.93) for the text with virtual manipulation group; M = 2.90 (SD = 1.95) for the narration with no manipulation group; and M = 3.88 (SD = 1.11) for the text with no manipulation group.

Sound Effects On Retention, Matching, And Transfer.

We predicted that realistic sound effects accompanying animation, such as the creaking of heavy wooden gates, the squeaking of valve handles, and the rush of water in and out of locks, would draw learners' attention to concurrent visual and verbal messages and thereby aid subjects in perceiving and retaining the sequence of steps, identifying pictured objects, and understanding and applying underlying principles of lock operations in novel situations.

Mean retention of the sound effects group was greater than that of the no sound effects group, although not significantly, F(1, 122) = 3.534, MSE = 3.645, p = .063, with M = 3.82 (SD = 1.92) for the sound effects group and M = 3.23 (SD = 1.89) for the no sound effects group. There was, however, a significant main effect for the sound effects group's gain on labeling pictured objects over that of the no sound effects group, F(1, 122) = 5.808, MSE = 2.458, p =.017, with M = 3.61 (SD = 1.51) for the sound effects group and M = 2.95 (SD = 1.69) for the no sound effects group. The sound effects group improved their transfer scores more than the no sound effects group, but this effect was non-significant, F(1, 122) = 2.401, MSE = 8.013, p = .124, with M = 5.20 (SD = 2.49) for the sound effects group and M = 4.50 (SD = 3.10) for the no sound effects group.

Summary

Overall, as measured by subjects' gains on the retention, matching, and transfer items, narration and text presentations appear equally effective as accompaniments to the self-paced animation. No manipulation appeared to be about as effective as virtual manipulation, and learners appeared to benefit to some degree from realistic sound effects synchronized to specific visual events. Virtual manipulation seemed to work better when subjects listened to narration, while non-interaction appeared to favor those who read text explanations. Posttest scores on retention and matching and transfer items were approaching the maximum score.

Discussion

Verbal Presentation

With respect to verbal presentation mode, in the case of self-paced multimedia instruction, we found no differences in this study between text and narration. On three different measures, subjects made similar gains regardless of verbal presentation mode.

This study builds on previous research (Mayer & Moreno, 1998) that showed a clear advantage for narration accompanying animation in system-controlled or timed multimedia instructional presentation. The present results may help to clarify the conditions in which narration or text is advantageous, based on the locus of control of the pace of instruction and, perhaps, the length and complexity of the text explanations. In our study, all learners controlled the pace of instruction to some extent, but only those who received text could control the pace of verbal instruction.

The split-attention effect (Chandler & Sweller, 1992), dual-processing model (Baddeley, 1992; Paivio, 1991) of working memory, and ideas about learner control (Kinzie, Sullivan, and Berdel, 1988) suggest an explanation. If two perceptual modalities, visual and aural, are engaged simultaneously, one with animation and the other with narration, then cognitive capacity is effectively increased and processing can proceed to the organizational stage of working memory (Mayer & Moreno, 1998). Similarly, if only the visual perceptual modality is used, but in a way that allows learners to attend to animation and text information at their own discretion and more ably deploy their limited cognitive resources, processing can also proceed to working memory.

Virtual Manipulation

These results do not shed much light on the utility of interactive multimedia as we have defined it here. Subjects who used their computer mouse to manipulate virtual screen objects such as opening and closing gates and valves and pulling the barge in and out of the lock before continuing to the next step, did not appear to learn more than subjects who merely watched the same animation and clicked a "Next" button to continue. This is particularly surprising with respect to the lack of difference in performance on the transfer questions, which attempt to measure how well learners grasped the causal model presented in the instruction. If in some way "causing" each step to take place in sequence did not help learners improve their grasp of the causal model, it is difficult to argue that virtual manipulation is worth the substantial effort and expense it takes to develop, at least in the context of learning causal systems.

Several explanations for the lack of significant results suggest themselves to the authors. First, it is possible that the virtual manipulation itself did have some effect, but some other, presumably closely related, design feature cancelled that effect. For example, it is possible that the virtual manipulation actually interfered with learning related to the outcomes we tested, as Gavora and Hannafin (1995) have cautioned. The time and attention expended in figuring out how to move the virtual objects may have increased the overall cognitive load, drew learners' attention away from the verbal explanation, and interfered with smooth processing in working memory. As Smith (2001) might put it, our interactive treatments may have forced learners to become more like pilots than co-pilots.

The interaction between presentation (narration vs. text) and virtual manipulation (vs. no manipulation) may support this explanation. Narration-listeners benefited more from virtual manipulation than did text-readers, suggesting that the visual perceptual mode used in conjunction with the tactile-kinesthetic perceptual mode involved in multimedia interactivity interferes more with words presented in the visual (text) mode than with the same words presented in the auditory (narration) mode. This would be consistent with assumptions of dual-processing, limited cognitive capacity, and split-attention.

Second, it is possible that virtual manipulation is of little or no value in assisting learners to increase their scores, for the particular learners and/or level and type of content we presented. For example, graduate students may be adept enough learners, generally, that virtual manipulation is of only marginal help. Posttest scores for all groups in our study were fairly high, indicating a possible ceiling effect. Another group of learners, fourth or fifth grade students, perhaps, who may rely more heavily on tactile-kinesthetic learning processes, might benefit more from this type of virtual manipulation, with this content.

Third, we might take a closer look at how we measured the learning that did take place. We presented the instruction with a combination of words, moving pictures, interactivity, and sound effects. We tested what subjects learned with verbal items only, except in the case of the matching items, which included a still image. Had we tested users' ability to correctly move a barge up or, in the case of transfer questions, down through a lock using the same degree of virtual manipulation, animation, and sound effects presented in the treatments, the results might have been different.

Sound Effects

Subjects who heard realistic sound effects associated with key objects and/or steps in the canal lock process showed more improvement than those who heard no sound effects, on items measuring retention of the sequence of steps, their ability to label pictures of key objects, and their ability to apply underlying principles of the process in new situations.

The only significant main effect was for matching labels to object pictures. The matching item was the only one that made use of an image, and this may have made it a better test of what subjects learned in the animated multimedia instruction, which relied extensively on similar imagery.

These results extend findings in previous studies (Calvert & Scott, 1989; Anderson & Levin, 1976) demonstrating that sound effects attract attention to concurrent visual and verbal information in television programs, thus facilitating perceptual selection and subsequent processing in working memory. The present study suggests that the same sort of perceptual selection may take place in computer-based multimedia instruction, in the context of explaining how causal systems work.

These findings are consistent with the proposition that sound effects provide redundant presentation of information with respect to the visual and verbal content about objects such as “sluice valves” or “swing beams” and actions such as “closing” or “opening.”

Instructional Design Implications

As Mayer and Moreno (1998) suggest, instructional designers should continually revisit important notions about needs assessment and content analysis for instructional multimedia design, and carefully think through learners’ cognitive demands as they design and build presentations. In particular, we can begin to make finer discriminations with respect to presentation modes, based on learner characteristics, content, and instructional context. Presentation mode clearly interacts with other design features, such as whether users will benefit more from timed, system-controlled or un-timed, self-paced instruction (see Mayer & Chandler, 2001).

The findings in this study on virtual manipulation suggest caution toward, but probably not abandonment of interactive multimedia as we have defined it here. The lack of significant differences in gains on scores between subjects manipulating virtual objects and those not manipulating objects, suggests that, at the very least, we look closely at proposed manipulative design elements and incorporate early and iterative user testing of such features during the design and development process, to insure we’re not throwing resources out an attractive window. On the other hand, the interaction between presentation and virtual manipulation suggest that, once we understand better how to design multimedia instruction as a whole, interactivity as we have defined it in this paper may be a useful tool in the instructional designer’s kit.

Another interesting area for experimentation uncovered in this study is in the use of instructional sound effects. By instructional sound effects we mean realistic sounds, synchronized with animation and/or verbal explanations, that draw attention to and underline concurrent verbal and/or visual information, and provide their own distinctive type of information about the character or quality of objects or actions involved in causal processes.

Limitations and Future Directions

The relative simplicity of the causal model and the instruction used in this study, with respect to the cognitive capacity and learning abilities of the graduate student subjects, may have limited the degree to which the various multimedia features facilitated learning. By contrast, the instruction Mayer and his colleagues (Mayer & Chandler, 2001; Mayer & Moreno, 1998) have used, on how lightning is formed, involves 16 steps, while the canal lock scenario we used consisted of only 5 steps. Posttest scores for all groups in our study were relatively high, indicating a possible ceiling effect. It would be useful to test some of the same multimedia features tested in the present study with a more complex causal model and instructional scenario, or with younger, less experienced learners.

Since it took text and narration groups, on average, about the same amount of time to complete their respective animated treatments, even though they could control the pace to a large extent in either condition, it would be interesting to examine whether simultaneous (animation-with-narration) or serial (animation-then-text) processing really takes about the same amount of time. If the animation-with-narration subjects perceived the visual and verbal information in roughly half the time as did the animation-then-text subjects, did the former take longer to organize the two channels in working memory that did the latter? That might explain the rough equivalence in treatment times for the two groups.

Are sound effects organized separately, or as part of the non-verbal channels of working memory? Might they interfere with (auditory) verbal information in perceptual processing, for example, and with images in the organizational stage of working memory? Or might they not interfere with working memory at all, but rather extend it, much as the visual and verbal channels appear to complement one another?

We hope to pursue some of these questions in future studies aimed at uncovering principles of educational multimedia design, based on understanding cognitive processing.

Educational technologists share some traditional beliefs about the effects of multimedia: that interaction and virtual manipulation are somehow equivalent to or at least promote active learning; that enhanced realism of visual and auditory representations goes hand in hand with increasing the authenticity of learning environments.

As multimedia capabilities become increasingly complex and expensive, we need to examine these assumptions more carefully.

The present study reminds us that the costs of specific educational multimedia features must be balanced against realistic assessment of the anticipated effects on learning that we can expect from those features. We need to learn more details about where and when to expend resources on specific features so that we avoid consuming the educational equivalent of "empty calories."

References

Alwitt, L., Anderson, D., & Lorch, E. (1980). Preschool children's visual attention to attributes of television. Human Communication Research 7(1), 52-67.

Anderson, D., & Levin, S. (1976). Young children's attention to "Sesame Street." Child Develoment 47, 806-811.

Baddeley, A. (1992, January 31). Working memory. Science, 255, 556-559.

Brooks, B.M., Attree, E.A., Rose, F.D., Clifford, B.R., & Leadbetter, A.G. (1999). The specificity of memory enhancement during interaction with a virtual environment. Memory 7(1), 65-78.

Calvert, S., & Scott, M. (1989). Sound effects for children's temporal integration of fast-paced television content. Journal of Broadcasting and Electronic Media 33(3), 233-246.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the desgin of instruction. British Journal of Educational Psychology, 62, 233-246.

Gavora, M.J., & Hannafin, M. (1995). Perspectives on the design of human-computer interactions: Issues and implications. Instructional Science 22, 455-477.

Kinzie, M.B., Sullivan, H.J., & Berdel, R.L. (1988). Learner control and achievement in science computer-assisted instruction. Journal of Educational Psychology, 80(3) 299-303.

Mayer, R., & Moreno, (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. Journal of Educational Psychology, 90(2), 312-320.

Mayer, R., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? Journal of Educational Psychology, 93, 390-397.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45(3), 255-287

Smith, G.G. (October, 2001). Interaction invokes reflection: Learning efficiency in spatial visualization. Interactive Multimedia Electronic Journal of Computer-Enhanced Learning. Retrieved March 14, 2002, from http://imej.wfu.edu/articles/2001/2/05/index.asp

Sweller, J. (1999). Instructional design in technical areas. Camberwell, Australia: ACER Press.

Acknowledgements

Chris Murphy, a graduate student and digital artist, created the original, photo-realistic 3D animation for the pilot study. Patricia Oliveira, also a graduate student and digital artist, drew, colored, and animated the simplified instructional animation sequence used in the main study. Thanks to 5th-grade teachers and students in the Lockport Unified School District who participated in the pilot study.

Building Instructional Design Expertise Through Metacognitive Scaffolding

Kathleen W. Ingram
Dawn E. Wright
University of South Alabama

Abstract

This study was an examination of the effectiveness a combination of instructional strategies for improving metacognitive thinking with the intent of moving novice instructional designers closer to "adaptive expertise" (Bransford, Brown, & Cocking, 2000). The participants, graduate students in an online beginning instructional design course, were divided into two groups that participated in one of two response-type instructional activities: responses to questions using a traditional question/answer format or responses to open-ended questions using a structured reflective journal. Each group also engaged in threaded discussion related to their specific questions as a way of creating meaning through shared knowledge and completed an individual instructional design project. The participants' level of reflective and metacognitive thinking was measured using a pre/post survey, answers to the response activities, and responses on the threaded discussion. The participants' change in instructional design expertise was measured with a pre/post case study questionnaire.

Expertise and Instructional Design

It is a well-documented fact that the information age has brought unprecedented change to business and industry. Because of these changes, the expectation from business and industry for the performance of today's graduates are much higher than for previous generations (Oblinger, 1998). The goal of instruction, in higher education, industry and the military, should be to help learners develop the ability to devise solutions to real-world problems by applying theoretical knowledge to their practice (Wong, Kember, Chung, Yan, 1994). Patel, Glaser, and Arocha (2000) state "… education needs to foster ways of learning that promote self-regulation and adaptability" (p. 258). How do we, as educators, help our students, both undergraduate and graduate, develop the skills that underlie expertise in their chosen profession?

Expertise is defined in many ways, depending on the attributes one associates with expert or outstanding behavior (Ericsson & Smith, 1991) and what those behaviors or actions contribute to one's profession (Kennedy, 1987). Earlier views of expertise cited technical skills, application of theory, critical analysis, and deliberate action as indications of expertise (Kennedy, 1987). This is somewhat exemplified in Ericsson and Smith (1991) statement that "… specific acquired characteristics underlie outstanding performance" (p. 4). However, Sternberg (1998) defines expertise as "…typically not an end state but is in a process of continual development" (p. 11). Bransford et al. (2000) use Hotano & Inagaki's (1986) term adaptive expertise to describe the kind of expertise that is flexible and adaptive to new situations. The definition this study chose is a combination of Ericsson, et al.'s acquired characteristics, Sternberg's (1998) developing expertise, and Bransford's "adaptive expertise".

Expertise is not an innate state but a learned and practiced level of ability Sternberg (1998). In fact, Ericsson and Smith (1991) state that there are certain general traits of experts in every field, and those traits are mostly acquired through practice over the course of many years.

Bransford, et al. (2000) describe six key principles, gathered from an extensive review of the research on expertise, that distinguish expert's knowledge from that of novice's knowledge. The principles, are that experts 1) notice features and meaningful patterns, 2) have a great deal of content knowledge that is well organized, 3) have contextualized knowledge, and 4) are able to retrieve important information with little cognitive effort, 5) may or may not be able to teach others, and 6) have varying levels of flexibility in their approaches to new situations. (Bransford, et al., 2000).

Most, if not all, of Bransford, et al.'s (2000) six principles of expertise are also found in the literature on instructional design expertise. For example, the following characteristics are discussed the literature on expert instructional designers: a deep and rich causal network of links of instructional design principles for evaluating and or deriving solution ideas, ability to use previous knowledge in problem framing and solution, develop breadth of problem, as well as rapid and efficient search of the problem and have excellent self-monitoring skills (e.g. LeMaistre, 1998; Perez & Emery, 1995; Roland, 1992). In fact, as evidenced by the International Board of Standards for Training Performance, and Instruction's (IBSTPI) Instructional Design Competencies (Richey, Fields & Foxon, 2001), an expert designer is expected to have both a multitude of domain specific skills, which require a deep and broad understanding of the discipline, as well as critical thinking and problem-solving skills.

The literature on expertise covers a wide range of disciplines (e.g. Chess, Teaching, Medicine, Sports, and Instructional Design). The purpose, however, of previous research is somewhat similar in its effort to determine 1) what characteristics or skills represent expertise in their particular field, 2) what general characteristics could be extrapolated across disciplines, and 3) comparison of novice to expert performance, and to a lesser extent 4) how someone becomes an expert (e.g. Bransford, et al., 2000; Ericsson & Smith, 1991; Kirschner, van Merriënboer, Sloep, & Carr, 2002; Le Maistre, 1998; Patel, et al., 2000; Perez, & Emery, 1995; Perez, Johnson & Emery, 1995; Snyder, 2000; Rowland, 1992). However, the focus of research in expertise has shifted from what expert behavior is to learning and instruction (Patel, et. al. 2000) for developing expertise.

Patel et al. (2000) state that there are three components to a theory of learning and instruction: 1) a theory of competent performance, exemplified by "domain experts", 2) a theory of how learning and development are acquired, and 3) a theory of about what strategies will best promote learning (p. 258). Therefore, this research study, rather than focusing on what instructional designers do or how people learn, attempts to address Patel, et al.'s third component by studying instructional, or learning, strategies for facilitating the attainment of expertise.

Metacognitive Development Through Reflective Thinking

Metacognition is "…the ability to reflect upon, understand, and control one's own learning" (Shraw and Dennison, 1994, p 460). Osman and Hannafin (1992) state "… metacognition refers to one's propositional, procedural, and conditional knowledge and the control of associated cognitive processes and activities" (p. 83). Metacognition, then, is an important component in developing not only discipline specific skills but also the general skills of flexibility and adaptation necessary for developing expertise.

Hacker (1998) states that most researchers identify two "…essential features of metacognition—self-appraisal and self-management of cognition" (Paris and Winograd, 1990, in Hacker, 1998). Self-appraisals are one's reflection on their knowledge and abilities Hacker, 1998

Metacognitive thinking can be learned. The implications for instructional design are to provide use "explicit strategies for younger versus older students and for novices versus experts, and use implicit, higher-order strategies for older learners and those with significant related prior knowledge" (Osman and Hannafin, 1992, p.94). Bransford, et al. (2000) state that "Teaching practices congruent with a metacognitive approach to learning include those that focus on sense-making, self-assessment, and reflection on what worked and what needs improving. These practices have been shown to increase the degree to which students transfer their learning to new settings and events" (p. 12).

Reflective thinking is a cognitive process that is believed to improve practice (Schön, 1983, 1987) through the development of higher-order thinking skills (Koszalka, Song & Grabowski, 2002), or provide a "…process leading from novice to expert in a field" (Salmon, 2002, p. 380). In fact Richey, et al. (2000) state that the participants who helped validate the competencies for the IBSTIPI competencies "…perceived generalized reflection as an essential element of successful design for all designers, novice and expert" (p. 72).

Ingram (2001) defines reflective thinking as "… an iterative *process* that allows practitioners to review their own practices by systematically considering the evidence- comparing their practices to the literature and their own experience, and generating an appropriate and relevant course of action" (p 27). The definition of reflection is somewhere between Sparks-Langer, et al. (1990), Schön's (1983, 1987) reflection-on-action and Dewey's (1933) "consecutive consideration", with an emphasis on reflection-on-action.

In an earlier study on reflective thinking training, Ingram (2001) found that while reflective thinking training, i.e., journaling combined with feedback and a reflective thinking seminar, helped the treatment group to reflect more at higher levels (either *framing* the problem and/or *generating solutions* within the current context and using language related to the theories and principles of that discipline) and at a "deep" level (a deep statement would not only describe the theory or principle but would try to give an explanation for why or how the solution is important in regard to learning) more often than the control group.

Based on the results of Ingram's 2001 study, the literature on expertise, and instructional design expertise, this study examined the use of several instructional strategies designed to increase students' metacognitive thinking. The strategies were a combination of reflective thinking, with included feedback and metacognitive scaffolding, and shared knowledge.

The Course Design

The course design was based on Russell, Reiser, Hruskocy and Ruckdeschel's (1999) description of a strategy for teaching project-based courses and the instructional materials used by Reiser (1999) in the a course for which the primary researcher served as a teaching assistant. The instructional strategies advocated by Russell, et al. included an authentic design project divided into smaller pieces to promote mastery, weekly reflective summaries to promote synthesis of content, modeling through sample projects, and an in-class project that allows students to practice and apply new knowledge and skills.

The instructional strategies specific to Reiser's course supported one or more of Reiser and Dick's (1996) six instructional activities for promoting learning, which are based on Gagné's (Gagné, 1985; Gagné, Briggs, & Wager, 1992) nine events of instruction. The six activities are: gaining attention (motivation), stating the objectives, recalling prior knowledge, presenting new information and examples, providing practice and feedback, and summarizing the lesson. These six activities correspond with the learning processes necessary to attend to, organize, and encode or relate new information to ideas and concepts already in the learner's memory so that the new information is more memorable (Driscoll, 2000).

However, the course was adapted for an online environment and to meet the needs of a different learner population. For example, to facilitate the group learning that occurs in a classroom, the students participated in a threaded discussion. The threaded discussion, similar to the sequencing of a traditional classroom based course, followed the individual reflective assignments.

In addition, to examine the effects of metacognitive thinking on instructional design expertise, a second type of "practice and feedback" instructional activity was added to the course. In his 1999 course Reiser used in-class design activities with whole-group feedback as a strategy for promoting practice and providing feedback. In the current study, a reflective journal, with embedded prompts (Lin, 2001), was used to scaffold the treatment groups' reflective thinking. In addition, each participant in the treatment group received feedback on their journals.

This study focused on how to help novice instructional designers develop adaptive expertise. Our initial research questions were: 1) How do we help instructional design students move from novice to expert, 2) Is metacognitive thinking a strategy that promotes expert decision-making, and 3) Will the instructional strategy of reflective thinking with feedback improve students' metacognitive thinking? Based on the literature, the researchers hypothesized that by providing metacognitive scaffolding novice instructional designers would increase their metacognitive skill level and that increase would enable them to solve more complex questions than those students who did not receive metacognitive scaffolding.

The use of a structured reflective thinking or metacognitive instructional strategy was hypothesized to increase students' metacognitive processes. In turn, students who have a higher level of metacognition were hypothesized to have a more holistic view of the instructional design process, and make design choices that are closer to an expert's choices

Methods

There were originally 18 participants in this study. The participants were both doctoral (3) and masters (15) students, ages 28-55, who were enrolled in an online beginning distance learning course. Due to normal attrition from the course, the final number of participants was 12, three doctoral and nine masters students. While the participants' academic instructional design experience varied, i.e. for some students this was their first course in their program and for others it was closer to the end of their program, most of them had limited professional instructional design experience.

This study employed a pre/post control group design, along with a qualitative component. At the beginning of the course the students completed two self-report surveys. The first survey, Shraw and Dennison's (1994) Metacognitive Awareness Inventory (MAI) consisted of 52 questions using a Likert scale to measure the participant's self-perception of their metacognitive abilities on two levels- knowledge of cognition, which was broken down into three components, and regulation of cognition, which was broken down into five components. Schraw and Dennison conducted several experiments to validate the MAI. The results from those studies showed that the multiple components had low internal consistency values. However, the results did show that there was a statistically significant relationship between knowledge and regulation of cognition ($r=.54$ and $.45$) and that the MAI does provide predictive information about performance (Shraw and Dennison, 1994, p. 471).

The second survey was designed to assess the participants' level of instructional design experience before the course began and after the completion of the course. The first part of the pre-survey asked for information on demographics and instructional design experience, while the second part was made of Ertmer and Quinn's (2003a) case study Lynn Dorman and questions from the Instructors manual (Ertmer & Quinn, 2003b). The post-survey was made of two case studies, Lynn Dorman was used as a comparison and a second case Denny Clifford (Ertmer & Quinn, 2003a) was included to account for pre/post testing threat as well as measuring the participant's ability to apply their instructional design knowledge to a new problem.

Based on the demographic data from the pre-survey and the data from the pre-MAI, the participants were matched by instructional design experience and metacognitive level. The two levels of design experience were High, i.e. the participant had either completed three or more courses in the Instructional Design program prior to this course or they had 2 courses and one year of instructional design experience, and Low, which meant they had no design experience and had completed less than three classes.

Participants were then placed into either a High or Low category of metacognitive awareness based on their mean scores on the eight subcategories of the MAI. In an effort to control for initial level of metacognitive awareness and design expertise the participants were matched and then randomly assigned to either the control or treatment group so that each group contained an equal number Low/Low, Low/High, High/High, and High Low participants.

The two levels of independent variable were the use of a metacognitive instructional strategy and the use of a traditional instructional strategy. While both groups received the same instructional materials, completed the same individual design assignments (situated learning) and participated in similar threaded discussions (shared meaning), the practice and feedback the two groups received were different. The control group completed traditional questions based on the chapter information and received corrective feedback based on their answers. The treatment group answered open-ended questions about the course content. For example in week six the control group was asked the following two questions as part of their assignment:

> Assume an instructional designer has written one objective that covers several substeps in an instructional analysis. Describe a problem that this decision might cause during later phases of the design process.
>
> The authors indicate that objectives help designers plan their instructional strategies and design their tests. Briefly (in one paragraph) describe how objectives can help designers perform one of these tasks. (Reiser, 1999)

The treatment group was asked the following questions for the same week:

> In addition to the concepts you wish to reflect on in your journal, please reflect on
>
> 1. How the parts of a performance object influence the rest of the instructional design process.
> 2. How the type of learning outcome (domain of learning) for a skill influences the way an objective is written and therefore the rest of the instructional design.
> 3. And, do you think the goal statement I had you write aligns with the textbook definition of a goal or more with that of a terminal objective? Why do you think I chose to have you do it one way over another?
>
> Please provide examples from experience or your own project in your discussion.
> Questioning the textbook author's assumptions/ procedures and/ or asking questions of
> yourself are also great ways to reflect at a deeper level!

The treatment group was required to use a journal that was structured to scaffold their reflective thinking in answering the weekly assignment questions. The journal was adapted from an earlier study, Ingram (2001). The earlier journal was developed to aid teaching assistants in higher education reflect on the use of a theory of learning as it related to their teaching practices. In the current study the journal's scaffolding questions were adapted to the context of instructional design practice. For example, in the earlier study one of the journal questions prompted the participants to "Focus on one or two of Reiser and Dick's instructional activities and explain why you think they were or were not effective based on the literature and… your personal teaching experience" (Ingram, 2001, p. 129) in an effort to scaffold the TAs' use of prior knowledge in their acquisition and application of new theories of learning. In the current study the same questions was stated as follows "Focus on two to three of the major concepts/terms and explain what you think they mean or describe how they are relevant or will be applied by using examples from the: Literature- use your book or readings to clarify your ideas Experiences- what past experiences (prior knowledge) can you bring in to help you understand/ apply these concepts."

While the treatment group received feedback to their journal entries in the form of prompts, the control group received corrective feedback. For example, the following prompt was in response to a journal entry in week three (reflection #2):

> Thad: In relation to instructional goals, the lecture brought forth the idea that it is possible that a goal may fall into more than one domain
>
> Prompt: "what do you think this means for the procedural analysis- diagram, and for the flow of instruction?"

An example of feedback given to the control group for the same week is stated below:

> Charlie: What I did notice was missing from the lecture model were the alternative connections between revising the instruction and the other steps in the process, such as development and selection of instructional material.

Prompt: Take another look- I believe they're connected through a dotted line?!?

The dependent variables were metacognitive thinking awareness and instructional design expertise. The first dependent variable, level of metacognitive awareness, was assessed with the following three measures: Shraw and Denninson's (1994) Metacognitive Assessment Inventory (MAI) as a pre/post survey, responses on the weekly threaded discussion, and answers to weekly assignments. The second dependent variable, level of instructional design expertise was measured with a pre/post case study analysis. The pretest used Ertmer and Quinn's (2001a) "Lynn Dorman", a case study that focuses on "the issues and typical obstacles involved in conducting a needs assessment" (Ertmer & Quinn 2003b, p. 43). The posttest included questions about the "Lynn Dorman" case study as well as questions about the "Denny Clifford" (Ertmer and Quinn, 2003a) case, which focuses on selecting the appropriate strategies for the learners, content, and context. The questions related to each case study were based on the questions in the instructors' manual.

Using the coding scheme in Table 1, the participants' journals were coded by the primary researcher using the qualitative data analysis program N6 (QSR International). A second coder analyzed several of the same journal entries and the two researchers used check coding to compare codes until they reached consensus on any discrepancies. Miles and Huberman (1994) state that "check coding not only aids definitional clarity but is also a good reliability check" (p. 64).

Table 1 *Inductive and deductive codes used in analysis of reflective journals and threaded discussions.*

Code	Operational Definitions
Experience- codes that emerged from the data	• ID Education –statement related to their own previous formal education in instructional design • ID Experience – statement related to their own previous professional experience designing instructional experiences • General experience –statement using their prior experience to explain or understand new knowledge.
Reflection- Ingram (2001)	"reflective thinking is an iterative *process* that allows practitioners to review their own practices by systematically considering the evidence – comparing their practices to the literature and their own experience, and generating an appropriate course of action (Ingram, 2001). 1. Description: provides description of problem (or triumph) related to instructional design but gives no rationale for problem. 2. Framing – provides explanation that includes terminology related to instructional design with rationale related to instructional design practice. 3. Solution generation – provides explanation with both rationale and solution based upon principles and theories related to instructional design.
Knowledge of Cognition- Shraw & Dennison (1994)	1. Declarative – knowledge about one's skills, intellectual resources and abilities as a learner. 2. Procedural – knowledge about how to implement learning procedures. 3. Conditional - knowledge about when and where to implement learning procedures (strategies).
Regulation of Cognition (Metacognition)- Shraw & Dennison (1994)	1. Planning – planning, goal setting, and allocation of resources prior to learning. 2. Information management – skills and strategy sequences used on-line to process information more efficiently (e.g. organizing, elaborating, summarizing, selective focusing). 3. Monitoring – assessment of one's learning or strategy use. 4. Debugging – strategies used to correct comprehension and performance errors. 5. Evaluation – analysis of performance and strategy effectiveness after a learning episode.
Expertise- Bransford, et. al (2000)	1. Pattern Recognition - notice features and meaningful patterns of information that are not noticed by novices 2. Content Knowledge a. Wide and Deep b. Organized – reflects deep understanding c. Contextualized – conditional d. Retrieval with little effort 3. Flexible approach to new situations (adaptive expertise)
Affective- emerged from data	Statement related to their affective state, e.g. "I feel…..," "Warming up to…..," "Been nervous about….."

An online pretest containing questions related to the case study Lynn Dorman (Ertmer & Quinn, 2003a) was administered in the first week of class. The questions were based on Ertmer and Quinn's (2003b) instructor manual. The pre and post surveys were coded using the expertise codes listed in Table 2. A similar process of check coding was used to develop the expertise codes and for coding the pre and post case study surveys.

Table 2 *Expertise scoring rubric for Pre/Post Survey**

Expertise
1. Pattern Recognition - notice features and meaningful patterns of information that are not noticed by novices (general)
i. Mental model or template of information they would need (Roland)
ii. Did not go back to problem-understanding phase (Roland)
iii. Search the problem space rapidly and efficiently (Le Maistre)
iv. Experts interpret, rather than just identify, the design problem (Perez & Emery)
v. Experts spend more time exploring the problem, i.e. developing problem breadth instead of solution breadth (Perez & Emery)
vi. Problem analysis- in-depth analysis, put off solutions until understanding of problem- although they did think of them early they didn't commit to them until considering their appropriateness for that problem (iterative process) (Roland)
vii. Perform extensive front-end-analysis (Le Maistre)
viii. In general experts were more reflective in their selection and use of a design strategy... i.e. they spent a lot more effort in the initial analysis & planning phase before selecting an instructional strategy (Perez & Emery)
2. Content Knowledge
a. Wide and Deep (general)
i. Linked information to experiences w/ similar problems (Roland)
ii. Continually challenged given information (Roland)
iii. Instructional & non-instructional solutions (Roland)
iv. Consider a wide range of factors in combination with one another (Perez & Emery)
b. Organized – reflects deep understanding
i. Problem representation- represented the problem as a deep and rich causal network of many links; used their knowledge of the principles of human learning and the performance that underlie the situation (Roland)
ii. Have a rich, well-organized knowledge base of instructional design. (Le Maistre)
iii. Represent problems at a deep level. (Le Maistre)
c. Contextualized – conditional
i. Problem interpretation- problem ill defined (yet simple); therefore solutions were tentative/ needed to be tested (Roland)
ii. Problem analysis- in-depth analysis, put off solutions until understanding of problem- although they did think of them early they didn't commit to them until considering their appropriateness for that problem (iterative process) (Roland)
iii. Hypothesized the results of interactions with individuals (Roland)
d. Retrieval with little effort
i. Search the problem space rapidly and efficiently (Le Maistre)
3. Flexible approach to new situations (adaptive expertise)
i. Switched to solution generation after understanding problem (Roland)
ii. Specified solution elements, proposing alternate solutions where material was unclear (to account for unknown conditions- needed more info from SMEs and customers (Roland)
iii. Have excellent self-monitoring skills (Le Maistre)

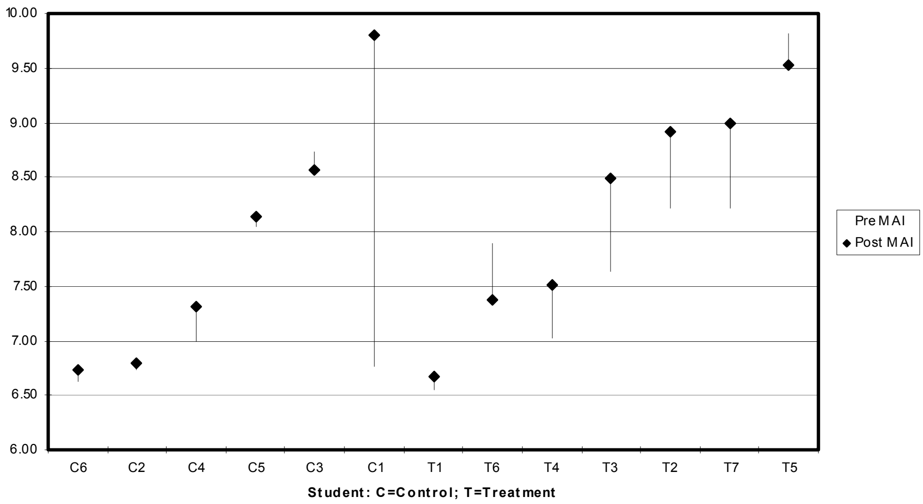
*Derived from Bransford, et al.'s (2000) General principles of expertise; Specific to Instructional Design Expertise: Le Maistre (1998), Perez & Emery (1998); and Roland (1992).

Preliminary Results

Metacognitive Awareness Inventory (MAI)

The MAI scores for each group are plotted in Figure 1, with a vertical line connecting the pre and post MAI scores for each student. Although the MAI Likert scale used for this study was 0 through 10, the vertical scale of Figure 1 has been truncated in order to show the changes in scores more readily, and to highlight the fact that none of the pre or post MAI scores were below 6.5. In general, the treatment group had larger changes in MAI scores than did the control group, although other than student C1, the changes in MAI scores of both groups are modest. Analysis of student C1's post MAI test indicated the student might have grown tired and selected the maximum point (10) on the Likert scale for essentially all of the 52 questions.

Figure 1, Change in MAI Scores



An ANOVA was conducted on the difference scores of the two groups (Student C1 was an outlier and was excluded from all quantitative tests). Levene's test for homogeneity of variances was significant, supporting the observation that the changes (variances) in the two groups' MAI scores were different. However, given the small sample size, and a significant Levene statistic, the Browne-Forsythe statistic was evaluated and found to not be significant, p= .34. Thus the mean change in MAI scores for the Treatment group, M = .306, SD = .55, was not significantly different from the mean change in scores for the Control group, M = .08, SD = .17.

Originally, we attempted to balance the two groups of students for pre-course MAI scores. However, attrition of students during the course resulted in potential confounding differences. A one-way analysis of covariance (ANCOVA) was conducted to account for these group differences in pre-course MAI scores. The independent variable included two levels: the Treatment group, received scaffolding feedback from the instructor on their reflective journals; the Control group, did not. The dependant variable was their post-course MAI scores and the covariate was their pre-course MAI scores. A preliminary analysis evaluating the homogeneity-of-slopes assumption indicated that the relationship between the covariate and the dependant variable did not differ significantly as a function of the independent variable, $F(1,8) = .002$, MSE = .22, $p = .97$, $\eta^2 < .001$.

The ANCOVA was not significant, $F(1,9) = 1.20$, MSE = .19, $p = .30$, $\eta^2 = .12$, indicating the null hypothesis that the adjusted mean post-course MAI scores of the two groups do not differ cannot be rejected. For these groups, the covariate, the pre-course MAI scores, is significant, $F(1,9) = 38.8$, $p < .001$, and accounts for 81% of the variance ($\eta^2 = .81$) in the mean post-course MAI scores of the two groups

The MAI consists of 52 questions divided into eight scales. Three of the scales, declarative knowledge (DK), Procedural Knowledge (PK), and Conditional Knowledge (CK), correspond to our operationalized definition of Cognition. The remaining five, planning (P), information management (IM), monitoring (M), debugging strategies (DS) and evaluation of learning (E) correspond to our operationalized definition of Metacognition. Table 3 contains the mean scores for the three MAI scales of Cognition and five MAI scales of Metacognition. It is interesting to note that while there is little apparent change in the pre and post scores for Cognition for both groups, the treatment group's Metacognition score increased but the control group's mean score was essentially unchanged. A paired-sample t test, however, indicated the metacognitive change for the treatment group was not statistically significant at the .05 level, $t(6)=-2.29$, $p = .067$. A one-sample t test of the combined MAI scores of the two groups indicates that the overall post-MAI score (M=7.93, SD=.94) is significantly different (p<.01) than the overall pre-Mai score (M=7.66, SD=.95).

Table 3 *Pre and Post Treatment MAI scores*

Group	N	Pre MAI Cognition		Post MAI Cognition	
		M	SD	M	SD
Treatment	7	8.12	1.04	8.13	1.43
Control	5	7.48	1.38	7.58	1.00
		Pre MAI Metacognition		Post MAI Metacognition	
		M	SD	M	SD
Treatment	7	7.83	1.11	8.27	.93
Control	5	7.42	.71	7.45	.79
Combined	12	7.66	.95	7.93	.94

Reflective Journal

Because of the incredible amount of qualitative data from the reflective assignments, threaded discussion, and pre/post survey, the data results reported in this paper are preliminary. The data reported here is from a comparison of two out of the 10 journal entries, the threaded discussions for both control and treatment group for the same weeks the journal entries were submitted, and six out of the 12 participants pre/post case-study surveys. Data from the control group' s assignments will not be included.

For this paper, the codes for the journal entries of week three, reflective entry #2, and week 12, reflective journal #10 were compared to examine possible changes in reflection or metacognition thinking. Table 2 is a comparison of the treatment group's combined codes for week six and week 12.

While Table 4 does not provide evidence of individual progress through the reflective process or how that relates to each participant's metacognition, it does show some interesting patterns. For example, the number of lower level reflective thinking entries, "Description", the metacognitive code "Planning" (see Table 1), and the code for "Prior knowledge- Prior experience" decreased dramatically from week two to week 10. In addition, while there was a noticeable increase in instances of "Reflection-framing" (see Table 1) from week six to week 10, there were only four instances of solution generation.

Table 4 *Combined frequency counts of codes for week two and twelve journal entries.*

Codes	Wk 2	Wk 10
Reflection-Description	110	32
Reflection-Framing	19	59
Reflection-solution generation	0	4
Cognitive-declarative	58	23
Cognitive-conditional	0	0
Cognitive-procedural	0	7
Metacognitive-planning	45	9
Metacognitive-info management	0	19
Metacognitive-evaluation	0	4
Metacognitive-debugging	0	2
Metacognitive-monitoring	0	0
Prior Knowledge-ID education	9	2
Prior Knowledge-ID Experience	7	0
Prior Knowledge-Prior Experience	50	2
Prior Knowledge-Experience-general knowledge	0	0
Prior Knowledge-PT experience	0	0
Affective	11	0

In our final results and conclusions we will provide examples of journal entries for each of these codes. In addition we will analyze the change in reflection and metacognition across all journal entries. We expect this data to provide evidence of how the journal scaffolds the participants' cognitive processes and if that scaffolding does help them to progress to a higher level of expertise than where they started.

Threaded Discussions

The qualitative analysis of the threaded discussions followed the same process and codes as used for the journals. Also similar to the journals, the two groups were given different prompts for the threads. We intended for the prompt for the treatment group to be less directive than the control group prompt and thus encourage more reflection. For example, treatment group's prompt in week three: "So, lets talk about the reflective process so far. What is working/ not working for you with the journal/ journal process and why? Feel free to help each other out with answers to questions and by providing ideas!" For week three, the control group was given this prompt: "So, lets talk about the FIRST & SECOND reflective assignment. Any thoughts or comments on the feedback you received on the first assignment? What were the pros/ cons of the second assignment? Feel free to respond to each other!"

It was our intent that the threaded discussions (threads) replace to the extent possible the classroom interaction between students rather than the interaction between students and the instructor. Although the literature indicates that online asynchronous discussions allow shy and unsure students more opportunity to participate, we found the threads were all too similar to the classroom with some students being very active and others participating minimally.

Coding of the threads is incomplete and continuing, but our preliminary analysis of the threads corresponding to the reflective journals weeks three or 12 indicates differences between two groups. As hypothesized, the treatment group's statements tended to be more reflective though some found the journaling process difficult. Trish: "I have been using the journal to write down key ideas. I have not been going into great detail on the journal. I have a hard time expressing what I know and comprehend through the written word. Maybe the journal will help me in this endeavor." But others were more comfortable: Toni: "I feel that the journal is a good source for me to apply prior knowledge. Also, when I write in the journal, I can see what I did wrong, and what I need to do. It also helps me to learn what was taught to me. I also can build each week from what I learned the prior week." Both of these statements were coded "knowledge of cognition, declarative." However the treatment group did not make any statements that were considered metacognitive or regulation of cognition in week 3 but did in week 12. Tracy: "I've come to realize that it is my lack of prior knowledge in this field impedes learning. I have also realized that I learn best with the face to face support of an instructor in a classroom." The control group did not make any statements that were coded as metacognitive in weeks three or 12.

Pre/Post Case-study Survey

The data in Table 4 only reports on eight out of the 12 participants results from pre to post survey, and the coding has not been completed to saturation. However, having stated that, there are some interesting results at this stage. One interesting result was that all participants in both the treatment and the control group had entries that were coded as "Novice/ solution" on both the pre and post surveys This code designates a statement of a solution without first stating the problem. Roland (1992) states that expert instructional designers put off solution generation until they understand the problem. For example in response to the post-survey question "What advise would you give Lynn about how to proceed with analysis and for possible recommendations?" one treatment group member responded:

Tom: I would recommend that Lynn meet with the management staff to disclose her findings and to offer possible solutions. These possible solutions would include training to bring employees up to date on the equipment and recommended safety procedures. I would also recommend a new safety program tied to incentives for attaining certain safety goals. Another recommendation might be to institute an incentive program tied to the new production rate. It appears that employees are expected to produce more without any monetary reward for achieving goals. I think if the employees were treated more fairly and management was willing to reward employees for achieving goals whether safety or production, the injury would be reduced dramatically.

A second interesting result was that on the posttest only two participants, Tom and Chris', responses were coded as exhibiting three out of four of the types of expertise (see Table 4). While each of these two participant exhibited different levels of expertise they both had "Pattern Recognition" and "Content knowledge/ Organized- Deep" (see Table 3) in common. In addition, these two participants were also the only ones to exhibit the level of expertise "Content knowledge/ Organized/Deep". An example of Tom and Chris' responses to the question "What information do you think Lynn needs to better understand the safety needs of these employees and the "gap" between the optimum and actual safety level?", were coded as "Content knowledge/ Organized/Deep", are listed below:

> Tom: I think the main information Lynn needs is input from employees and how they feel about the safety issues and why the don't' use the safety equipment supplied to them. By talking with the employees Lynn can determine if employees actually need training about safety or if there are other issues such as motivation or incentive that cause them to disregard the rules set (by) management. Lynn also needs information about the upper management staff and their perceptions of the problems and what solutions they recommend. If Lynn discovers that the increased production speed is the main cause of accidents and upper management requires speed she may have a problem that will be difficult to solve. (note: this same passage

was also coded as "Pattern recognition" and "Content knowledge/ Large amount")

Chris: Though a 'zero' accident rate would be ideal, it is probably not realistic. As such, she (Lynn) would need to obtain historical accident rate data for the company and also accident rate data for comparable companies, both in size and industry. In determining the gap, she will need to establish the 'optimum'- be it the industry average or better. Additionally, she will need to determine the root cause of recent accidents- was it PPE (personal protective equipment), production rate, equipment etc. Without knowing the cause, it will be difficult to address the 'gap'. (note: this same passage was coded as "Pattern recognition").

After the five remaining participants', three from the treatment group and two from the control group, pre and post survey's have been coded we will compare the results both across and between groups. From the preliminary data we believe that an analysis of which codes are most often exhibited by which participants as well as what codes are used for the same passages will provide some evidence for possible inter-related cognitive processes when moving from novice to expert.

Table 4 *Pre/Post Case Studies frequency counts by # of participants per group with codes.*

	Treatment		Control	
Novice:	Pretest	Posttest	Pretest	Posttest
Cognition/declarative	1/4	2/4	1/4	0
Novice/Identify/Not analyze Problem	2/4	3/4	3/4	4/4
Novice/Little depth	4/4	4/4	3/4	3/4
Novice/Solution	4/4	4/4	4/4	4/4
Novice/Lack of pattern recognition	3/4	3/4	1/4	2/4
Expertise:				
Pattern recognition	2/4	2/4	1/4	3/4
Content knowledge/Large amount	1/4	1/4	1/4	0
Content knowledge/Contextualized	1/4	0	1/4	2/4
Content knowledge/ Organized/Deep	0	1/4	0	1/4

Preliminary Conclusions

Obviously incomplete data does not provide strong evidence for any conclusions drawn from that data. After completing the analysis of data from the reflective journals, the traditional question/answer assignments, the threaded discussions, and the pre post surveys we will be better able to provide strong conclusions. In that analysis and the subsequent conclusions we will examine the effects of the process of using a journal to scaffold reflection and metacognition on the treatment group's threaded discussion, MAI scores and post test.

References

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (2000). *How people learn: Brain, mind, experience and school*. Washington, DC: National Academy Press.

Dick, W., Carey, L., & Carey, J.O. (2001). The Systematic Design of Instruction (5th ed.). New York: Longman.

Driscoll, M. P. (2000). *Psychology of learning for instruction* (2nd ed.). Boston, MA: Allyn and Bacon.

Ericsson, K. A. & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In

Ericsson, K. A. & Smith, J (Eds.) *Toward a general theory of expertise: Prospects and limits. Cambridge*,
NY: Cambridge University Press.

Ertmer, P. A. & Quinn, J. (2003a). *The ID casebook: Case studies in instructional design* (2nd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Ertmer, P. A. & Quinn, J (2003b). *Instructor's manual to accompany the ID casebook: Case studies in instructional design* (2nd ed.. Upper Saddle River, NJ: Merrill Prentice Hall.

Gagné, R. M. (1985). The conditions of learning: And the theory of instruction. (4th ed)

Gagné, R. M., Briggs, L. J. & Wager, W. W. (1992). *Principles of instructional design* (4th ed). Orlando, FL: Harcourt Brace Jovanovich, Publishers.

Hacker, D. J. (1998). Definitions and empirical foundations. In Hacker, D. J., Dunlosky, J. & Graesser, A. C. (Eds.)*Metacognition in Educational Theory and Practice*, (pp. 1-23). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Ingram, K. W. (2001). *The effects of reflective thinking training on TAs' reflective thinking, use of instructional activities, instructional effectiveness, motivation to teach and their students' attitudes toward instruction*. Unpublished doctoral dissertation, Florida State University, Tallahassee

Koszalka, T. A., Song, H, & Grabowski, B. (2002). *Examining learning environment design issues for promoting reflective thinking in web-based PBL*. Paper presented at the 2002 American Research Association, New Orleans, LA.

Kennedy, M. M. (1987). Inexact sciences: Professional education and the development of expertise. In Rothkopf, E.Z. (Eds.) (pp 133-167). *Review of the Research*. Washington, DC: American Educational Research Association.

Kirschner, P., van Merriënboer, J., Sloep, P, & Carr, C. (2002). How expert designers design. *Performance Improvement Quarterly*, 15 (4), 86-104.

LeMaistre, C. (1998). What is an expert instructional designer? Evidence of expert performance during formative evaluation. Educational Technology Research & Development 46, 3, 21-36.

Lin, X. (2001). Designing metacognitive activities. *Educational technology, Research & Design*, 49 (2), 23-40.

Miles, M. B. & Huberman, A. M. (1994). An expanded sourcebook: Qualitative data analysis (2nd ed). Thousand Oaks, CA: Sage Publication, Ltd.

Oblinger, D. G. & Verville, A. (1998). *What business wants from higher education*. Phoenix AZ: The Oryx Press.

Osman, M. E. & Hannafin, M. J. (1992). Metacognition research and theory: Analysis and implications for instructional design. *Educational Technology, Research & Development*, v40 (2), 83-99.

N6 (2002). N6 [Computer software]. Melbourne, Australia: QSR International Pty. Ltd.

Patel, V. L.& Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In Ericsson, K. A. & Smith, J (Eds.). *Toward a general theory of expertise: Prospects and limits*. Cambridge, NY Cambridge University Press.

Patel, V. L., Glaser, R. & Arocha, J. F. (2000, February). *Cognition and expertise: acquisition of medical competence*. Paper presented at workshop held in Sidney, BC.

Perez, R. S. & Emery, C. D. (1995). Designer thinking: How Novices and experts think about instructional design. *Performance Improvement Quarterly*, 8 (3), 80-95.

Perez, R. S., Johnson, J. F. & Emery, C. D. (1995). Instructional design expertise: A cognitive model of design. *Instructional Science*, 23, 321-349.

Reiser, R. (1999). Selected instructional materials for EME 5603 Introduction to Systematic Instructional Design.Florida State University.

Reiser, R. and Dick, W. (1996). *Instructional planning: A guide for teachers* (2nd ed.). Needham Height, MA: Allyn and Bacon.

Richey, R. C., Fields, D. C., & Foxon, M. (2000). *Instructional design competencies: The standards*. : Syracuse, NY: ERIC Clearinghouse on Information Technology.

Rowland, G. (1992). What do instructional designers actually do? An initial investigation of expert practice *Performance Improvement Quarterly*, 5 (2), 65-86.

Russell, J. D., Reiser, R. A., Hruskocy, C., & Ruckdeschel, C. (1999). Strategies for teaching project-based courses. *Educational Technology* (March-April) 56-59.

Salmon, G. (2002). Mirror, mirror, on my screen: Exploring online reflections**.** *British Journal of Educational Technology*, 33 (4), 379-392.

Schraw, G. & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475.

Schön, D. A. (1983). *The reflective practitioner: How professions think in action*. United States, USA: Basic Books, Inc.

Schön, D. A. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass Publishers.

Snyder, J. L. (2000). An investigation of the knowledge structures of experts, intermediates an novices in physics. *International Journal of Science Education*, 22 (9) 979-992.

Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher*, 27 (3) 11-20.

Wong, F. K. Y., Kember, D., Chung, L. Y. F, Yan, L (1995). Assessing the level of student reflection from reflective journals. *Journal of Advanced Nursing*, 22, 48-57.

Power Relations Among Students in Online Discussions in Adult and Higher Education

JuSung Jun
Thomas Valentine
The University of Georgia

Abstract

The purpose of this paper is to present the findings of a study dealing with the nature of power relations in two online graduate courses. The findings are that (1) There were no statistically significant differences of power inequality in gender and race among adult students in terms of powerful/powerless language use, and (2) the female and the Caucasian groups that were majorities in the current study had higher mean ranks through all five indicators of powerful and powerless language use than those of the male and the African-American groups.

Introduction

Power is a key element in all human interactions. Tisdell (1993, 203) points to the structural inequality that exists in society, the power disparity between racial minorities and the white majority, between the poor and the wealthy, the undereducated and the educated, and women and men and how these power relations are reproduced and maintained through the educational process. Cunningham (2000, 573) argues that much of the field of adult education's rhetoric centers on the learners, as if the learners are disembodied creatures and as if the social context, the social structures, the social class in which we all exist do not affect the process of education. Wilson and Cervero (2001), in citing Livingston (1983), contend that to practically confront the world of inequity, we need to understand the way it is, have a vision for what it should be, and have strategies for achieving our vision. Brookfield (2000) notes that power is ubiquitous in adult classrooms, inscribed in the practices and process that define the field. He maintains that when we become aware of the pervasiveness of power we begin to notice the oppressive dimensions to experiential practices that we had thought were neutral or even benevolent (40). Discussion is usually considered as a powerful tool for the development of pedagogic skills such as critical thinking, collaboration, and reflection as well as for the improvement of democratic communication. Based on his experience as a learner or a facilitator in a discussion group, Brookfield (2001) underscores that unless adult educators create a space for those voices that would otherwise be excluded by default, discussion reproduces structures of inequity based on race, class, and gender that exist in the wider society. As Wilson and Cervero (2001) point out, the systems of power that structure all action in the world are an inescapable facet of social reality and usually asymmetrical in that they privilege some people and disadvantage others. The pressing problem is to disclose the unequal power relations between people who have structural privilege and those who do not and then work to reduce this inequality.

Although there is a body of literature that discusses the types of interaction or the factors influencing interaction in online discussions for adult learners, there has been a lack of research that specifically examines the nature of power relations among adult learners in online discussions. The present study employed critical discourse analysis as a methodological approach in an attempt to understand the nature of power relationships that occurred in the written discourse of two online graduate courses.

The purpose of this paper is to present the findings of a study dealing with the nature of power relationships in a specific online learning setting. This purpose is guided by the following research questions: (1) To what extent are power and powerlessness present in an online classroom? (2) To what extent can manifestations of power and powerlessness be explained by personal characteristics of students?

Power Relations in Discussions

There are a considerable number of studies describing online discussions or comparing online and face-to-face discussions in adult and higher education. These studies have mostly focused on such subjects as (1) the comparative advantages and disadvantages of online and face-to-face discussions, (2) the individual and group interaction patterns in online discussions, (3) the development a methodology for analyzing transcripts, and (4) the relative merits of various instructional strategies for online discussions. However, most of these studies have given little attention to the nature of power relationships and the ways in which power and privilege are manifested in the online discussions.

Tisdell (1993) examined how power relationships, primarily based on gender but including race, class, and age, were manifested in higher educational classrooms of adult students through class observations, interviews, and document analysis. She observed several significant facts in terms of power relations: (1) the students who benefited from more interlocking systems of structural privilege tended to have more power and a more dominant role in the classroom than their less privileged peers, (2) the students contributed to reproducing structured power relations in their reification of patriarchal values, (3) a male professor tended to exert more control than a female professor, and (4) the middle-aged women with more education tend to be more participatory, at least in classes where affective forms of knowledge are valued.

Grob, Meyers, and Schuh (1997) examined sex differences in powerful/powerless language such as interruptions, disclaimers, hedges, and tag questions in the small group context of a higher education classroom by juxtaposing two competing theoretical frameworks: "dual cultures" and "gender similarities." Their findings revealed that there were no significant differences between women and men in their use of interruptions, hedges, and tag questions, which supports "gender similarities" approach to understanding sex differences and not the dominant "dual cultures" approach for investigating sex differences. In other word, there was no evidence that men use more powerful language while women use powerless language.

Carli (1990) observed mixed-sex and same-sex dyads consisted of undergraduate students to examine effects of gender composition on language and of language on gender differences in influence. She found that (1) women were more tentative than men, but only in mixed-sex dyads, (2) women who spoke tentatively were more influential with men and less influential with women, (3) women in same-sex dyads were more likely to use intensifiers and verbal reinforcements than men, whereas no gender differences emerged in mixed-sex dyads, and (4) men were equally influential, whether they speak assertively or tentatively.

Some studies focused on gender differences in communication patterns in online discussions. McAllister and Ting (2001) explored gender differences in computer-mediated communication in web-based college courses, analyzing the 456 discussion postings of 34 students in two online college courses. Each discussion posting was analyzed for seven variables: total comments, comment length, readability level, intended audience, purpose, references, and format. The findings of the study suggested that male and female discussion items differed significantly in length, use of indicators to specify a particular reader, purpose, and use of formal signature. However, male and female discussion items did not differ in frequency, readability, intended audience, or references to personal experience or outside sources.

In a similar vein, Fahy (2002) investigated gender-related communication differences in the use of linguistic qualifiers (e.g., I think, may/might, often, perhaps, etc.) and intensifiers (e.g., always, certainly, of course, only, very, etc.) in a computer conference transcript by examining a transcript consisting of 356 student postings. The result s of the study suggested a tendency for women to use more of the forms thought likely to sustain dialogue (qualifiers, conditional and parenthetic statements, and personal pronouns), while men's postings generally contained fewer qualifiers and more intensifiers.

Although all of this literature contributes to our understanding of power in online learning, it is clear that additional work is needed if we are to understand power dynamics in this rapidly growing educational format. This study explored the ways in which power and privilege are expressed in the online discussions in higher education. Our methodology is similar to that used by Fahy (2002), whose focus on gender-based differences in using qualifiers and intensifiers anticipates our focus on gender and race differences in the use of powerful/powerless languages.

Research Participants and Data Collection

The two online classes, "A" and "B", selected for this study were Master's level courses in a professional school at a large state university in the southeastern United States in Spring 2002. The same instructor taught the two online classes. In addition, the web-sites, instructional materials, and learning activities for the two courses were identical. The graduate students enrolled in the two classes consisted of 10 males and 31 females. All but two of the students were part-time students who had full-time jobs. Twenty-nine of the students were white, ten were African American, and one was of unknown race. Although precise age data were unavailable, it is known that the students' ages ranged from mid-20s to mid 50s. The students knew each other based on their short biography submitted before the class began; some students had included photographs with their biographies, and these were posted.

The online course contained 10 units related to the concepts of educational research. Each class employed 3 mixed-sex small group discussions in each of the 10 learning units. More specifically, for each learning unit, the instructor gave discussion questions or set a discussion task. Each member of the group was required to make at least two "substantive contributions" to the discussion of each unit. The instructor didn't participate in the small group discussions, believing that his participation could negatively affect the level of student problem-solving. However, in two of the 10 units, content area experts joined the online discussions.

Substantive contributions were defined as having three major characteristics: (1) The contribution must relate either to the discussion task the instructor set or to the comments made by other group members, (2) it must be well thought-out and well crafted, and (3) it must be at least a two sentences in length. Each discussion was time-bound; there was a tightly controlled time period during which students must make their contributions. The total span of the discussion activity during the first through eighth unit was 1 week; for the final two sessions, the discussion was 2 weeks in duration. Ultimately, there were a total of 1340 postings made in the two classes over the span of the semester.

Data Coding and Analysis

In this study discussion postings were analyzed and coded based on techniques of critical discourse analysis, closely keeping in mind the primary tenets of critical discourse analysis (CDA). As Fairclough (1995, 6) notes that there is no set procedure for doing discourse analysis, people approach it in different ways according to the specific nature of the project, as well as their own views of discourse. Van Dijk (1993, 235) also points that critical discourse analysis is far from easy. . . . it requires true multidisciplinarity, and an account of intricate relationships between text, talk, social cognition, power, society and culture. Joyce (2001) stresses, by taking a position, researchers must be self-reflexive in terms of their interpretations and analyses and maintain some distance in order to avoid producing analyses that map directly onto their own personal beliefs.

In this study, our analysis was based on indicators of power/powerlessness drawn from studies by Carli (1990), Grob, Meyers, and Schuh (1997), Jun and Park (2003), Kollock, Blumstein, and Schwartz (1985), and Tisdell (1993) to identify and analysis power relations among participants. These indicators are total words, postings, and acknowledgement by others for powerful languages and disclaimers, hedges, and tag questions for powerless languages.

Powerful Languages: Total words, Postings, and Acknowledgement by others

Dovidio, Ellyson, Keating, Heltman, and Brown (1988) found that high status or high dominance men and women display a greater amount of verbal and nonverbal power, as measured by the amount that subjects look at their partners while speaking and look away while listening (as cited in Carli, 1990). Tisdell (1993) points out that the students who benefited from more interlocking systems of structural privilege tended to have more power, playing the dominant role presented by outspokenness in the class. In the current study, we assume that a person who has more words (or postings) is more powerful than those who have fewer words (or postings).

Acknowledgement by others can be defined as the number of postings which receive responses from others. A person who has more acknowledgements by others is more powerful than those who have fewer acknowledgements by others (Jun and Park, 2003).

Powerless Languages: Disclaimers, Tag-questions, and Hedges

Disclaimers are expressions of uncertainty (e.g. "I guess," "I suppose," "I don't know much but," "I'm not an expert but"), indicating a lack of power (Carli, 1990).

Tag-questions are shortened questions added to a declarative sentence. Grob, Meyers, and Schuh (1997) notes that tag-questions are considered to be forms of powerless speech because they turn a declarative statement into a question, making the speaker appear more uncertain and less assertive. Carli (1990) found that women used more tag-questions than men in both same-sex and mixed-sex dyads. Kollock, Blumstein, and Schwartz (1985) maintain that the more powerful person of either sex is more likely to interrupt and to be more successful at it, whereas the less powerful person tends to use more tag questions.

Hedges can be defined as adverbs or adverb phrases which contain little or no meaning, convey either moderation or no particular meaning at all (Grob, Meyers, and Schuh, 1997). Carli (1990) found that women used more hedges than men in both same-sex and mixed-sex dyads.

Our analysis consisted of examining the 1340 postings for the two classes-all of which were converted into (PDF) electronic files. We accomplished coding on a passage-by-passage basis. In addition, we did not feel in need of checking reliability because the powerful/powerless languages of interests were relatively unambiguous and we followed the words/phrases that the previous studies found. Table 1 presents a list of words/phrases coded in three categories which are the indicators of the use of powerless languages.

Table 1. *List of All Words/Phrases coded as Powerless Languages*

Disclaimers (N = 518)	Tag-Questions (N = 48)	Hedges (N = 1664)
I am a lot lost!	Am I really still foggy?	A bit
I'm (was) confused	Are we correct in how we understand this?	A little (bit)
I'm not sure/unsure	Doesn't it?	A tad
I'm stuck here	Huh?	About
I (would) assume	Is that correct?	Almost
I consider	Is that right?	Anything like that
I don' t (do not) know	Is this common?	Around
I don't see	Is this how you see it?	Could be
I don't understand	Is this true?	(Please) help!
I (also/do/just/still/want to/would) feel	Isn't it?	(In my) humble opinion
I (would) guess	Make sense?	Just my thought/a thought
I have no idea	So, what is the correct definition?	Kind of/sort of
I mean	(Am I) right?	Kinda
I suspect	Would that work?	Like
I want to make sure I understand this properly	(Am I) wrong?	(Un)likely
I wonder (am wondering)		Look like
It (does/ would) seems		May (not)
		Maybe
		Might (not)
		More toward
		Most(ly)
		Otherwise
		Perhaps
		Pretty (much)
		Probably
		(Sometimes) seem (like)
		Seemingly
		Silly
		Slightly
		Something like (that)
		Somewhat
		Sound like
		That much
		Usually
		Whatever
		Would be
		You know

To test the research question, we conducted frequency analysis and the Mann-Whitney U test using SPSS 11.0, using gender and race as the independent variable. Descriptive statistics were generated for each variable: means, standard deviation, standard error, and min and max values.

Findings

Findings Related to the Research Question # 1

Table 2 shows the items which are over 5% in each powerless language category. The principle findings were:

- In the category of disclaimers, "It (does/ would) seems" accounts for the most part of the total frequencies with 222 (42.86%) out of a total of 518 observations, followed by "I (would) guess" with 71 (13.71%) observations. "I (also/do/just/still/want to/would) feel" placed third with 66 (12.74%) the observations. The others account for almost 23% of total percentage with 118 frequencies.
- In the category of tag-questions, "(Am I) right?" was the most frequently used, with 18 (37.50%) observations, followed by "Make sense?" with 13 (27.08%) observations. The others account for 35.42% of total percentage with 17 frequencies.
- Finally, in the category of hedges, "May (not)" held on to its No. 1 spot among most-used hedges with 417 (25.06%) observations. "Might (not)," "Probably," "Maybe" are hanging on to their respective rankings as the No. 2, 3 and 4 most-used hedges with 255 (15.32%), 101 (6.07%), and 99 (5.95%) observations respectively. The others account for 47.60% of total percentage with 792 frequencies.

Table 2. *Specific Indicators of Powerless Languages in Online Discussions*

Indicators	Frequency	Percentage
Disclaimers		
It (does/ would) seems	222	42.86
I (would) guess	71	13.71
I (also/do/just/still/ want to/ would) feel	66	12.74
I'm not sure/unsure	41	7.91
Others	118	22.78
Total	518	100.00
Tag-Questions		
(Am I) right?	18	37.50
Make sense?	13	27.08
Others	17	35.42
Total	48	100.00
Hedges		
May (not)	417	25.06
Might (not)	255	15.32
Probably	101	6.07
Maybe	99	5.95
Others	792	47.60
Total	1664	100.00

Table 3 and 4 show the descriptive statistics (means, standard deviations, minimum and maximum values, skewness, and kurtosis) for each variable according to gender and race groups.

Table 3. *Powerful/Powerless Language Use for Male and Female Groups*

Group	Variables	M	SD	Min	Max	Skewness	Kurtosis
Male (N = 10)	Total Words	3373.40	1085.373	1883	5035	.520	-1.128
	Postings	28.80	5.203	21	36	-.201	-1.530
	Acknowledgement by Others	9.00	5.375	1	20	.725	.935
	Disclaimers	10.40	5.700	1	19	-.279	-.492
	Tag-Questions	1.10	1.449	0	3	.608	-1.937
	Hedges	37.30	15.019	15	59	-.141	-.731
Female (N = 31)	Total Words	3699.10	1281.866	1622	7573	1.035	1.553
	Postings	33.94	12.757	15	73	1.151	1.640
	Acknowledgement by Others	10.39	6.551	0	27	.649	.559
	Disclaimers	13.68	7.743	2	31	.423	-.824
	Tag-Questions	1.19	1.579	0	6	1.721	2.537
	Hedges	41.65	18.340	12	88	.492	-.131

Table 4. *Powerful/Powerless Language Use for African-American and Caucasian Groups*

Group	Variables	M	SD	Min	Max	Skewness	Kurtosis
African-American (N = 11)	Total Words	3114.64	867.664	1883	4613	.200	-.859
	Postings	29.73	10.355	19	49	1.012	-.071
	Acknowledgement by Others	6.55	4.547	0	12	-.052	-1.650
	Disclaimers	8.55	5.628	1	20	.828	.515
	Tag-Questions	.73	1.489	0	5	2.788	8.325
	Hedges	34.73	15.080	16	60	.536	-.944
Caucasian (N = 29)	Total Words	3783.17	1321.916	1622	7573	.944	.977
	Postings	33.55	12.073	15	73	1.556	3.261
	Acknowledgement by Others	11.00	6.193	0	27	.758	.878
	Disclaimers	14.38	7.481	3	31	.342	-.658
	Tag-Questions	1.38	1.545	0	6	1.242	1.402
	Hedges	41.90	17.793	12	88	.458	.339

Findings Related to the Research Question # 2

As seen in Table 5, the results of the Mann-Whitney U test show that there were no statistically significant differences in the use of powerful/powerless languages between male and female groups at a significance level of .05.

Table 5. *Differences between Male and Female Groups on the Use of Powerful/Powerless Languages*

	Male (n =10)		*Female (n = 31)*				
Variables	Mean Rank	Sum of Ranks	Mean Rank	Sum of Ranks	Mann-Whitney U	Z	p
Total Words	*18.90*	*189.00*	*21.68*	*672.00*	*134.00*	*-.638*	*.524*
Postings	*17.70*	*177.00*	*22.06*	*684.00*	*122.00*	*-1.004*	*.316*
Acknowledgement by Others	*18.90*	*189.00*	*21.68*	*672.00*	*134.00*	*-.640*	*.522*
Disclaimers	*17.40*	*174.00*	*22.16*	*687.00*	*119.00*	*-1.095*	*.273*
Tag-Questions	*20.00*	*200.00*	*21.32*	*661.00*	*145.00*	*-.323*	*.747*
Hedges	*19.30*	*193.00*	*21.55*	*668.00*	*138.00*	*-.516*	*.606*

The Mann-Whitney U test (see results in Table 6) shows that there were significant differences in the use of powerful/powerless languages between African-American and Caucasian Groups. Specifically, there are significant differences in "acknowledgement by others" ($Z=-2.069$, $p=.039$) and "disclaimers" ($Z=-2.125$, $p=.034$) between the two racial groups at a significance level of .05. These results mean that the Caucasian group had more "acknowledgement by others," which is an indicator for the use of powerful language, than the male group. At the same time, interestingly, this group used more "disclaimers," which is an indicator for the use of powerless language, than the other group. The standardized mean differences between the two samples equaled .77 and .83 respectively, large effect sizes. This suggests that the differences in "acknowledgement by others" and "disclaimers" between the two groups are statistically significant as well as substantively significant.

Table 6. *Differences between African-American and Caucasian Groups on the Use of Powerful/Powerless Languages*

Variables	*African-American (n =11)* Mean Rank	Sum of Ranks	*Caucasian (n = 29)* Mean Rank	Sum of Ranks	Mann-Whitney U	Z	p
Total Words	*16.09*	*177.00*	*22.17*	*643.00*	*111.00*	*-1.469*	*.142*
Postings	*17.23*	*189.50*	*21.74*	*630.50*	*123.50*	*-1.092*	*.275*
Acknowledgement by Others	*14.32*	*157.50*	*22.84*	*662.50*	*91.50*	*-2.069*	*.039**
Disclaimers	*14.14*	*155.50*	*22.91*	*664.50*	*89.50*	*-2.125*	*.034**
Tag-Questions	*16.00*	*176.00*	*22.21*	*644.00*	*110.00*	*-1.588*	*.112*
Hedges	*17.05*	*187.50*	*21.81*	*632.50*	*121.50*	*-1.152*	*.249*

Note: * Correlation is significant at the .05 level.

Discussion

The results of the study show that there were no statistically significant differences in the use of powerful/powerless languages based on gender. This suggests either that there are no notable differences with respect to how men and women use powerful/powerless language in online discussions—at least so far as such use is captured by the variables used in this study.

With respect to race, the results revealed significant differences in "acknowledgement by others" and "disclaimers" between the African-American and the Caucasian groups (Z=-2.069, p=.039; Z=-2.125, p=.034) at the significance level of .05. The Caucasian group received more "acknowledgements by others" which is an indicator of powerful language, yet used more "disclaimers" which is an indicator of powerless language. These results appear to be in conflict, which presents a very real challenge to the validity of the measures used in this study. A reexamination of both the conceptualization and the operationalization of these variables are currently underway.

If the measures used are, in fact, valid, then our findings suggest the possibility that the online discussion environment attenuates the power of gender-based privilege and, perhaps, undercuts race privilege—at least with respect to the use of disclaimers. Both researchers and practitioners need to embrace the possibility that online learning contexts might fundamentally alter power dynamics of discussions by eliminating the impact of physical appearance, size, body language, and tone of voice. The two-dimensional, linear, asynchronous nature of online discussion offers a very real contrast to the three-dimensional, sometimes-chaotic world of face-to-face discussion.

As mentioned earlier, unless adult educators create a space for those voices that would otherwise be excluded by default, discussion will reproduce structural inequity based on race, class, and gender existing in the wider society (Brookfield, 2001). Brookfield (2001) notes:

> The adult discussion leader cannot be a laissez-faire facilitator, exercising a minimum of control. Taking this stance only serves to allow patterns of inequity present in the wider society to reproduce themselves automatically in the classroom. Instead, the teacher must intervene to introduce a variety of practices to insure some sort of equity of participation (221-22).

Adult education is a site for the struggle for knowledge and power (Cervero and Wilson, 2001, 3), and adult educators must pay attention to power inequality that might exist in the classes they teach to ensure both democratic participation and high quality learning. Facilitators of the online courses need to monitor participation patterns and intervene as necessary to ensure inclusive dialogue and reduced marginality. To do this, Brookfield (2001, 222-223) suggests three steps:

- First, make sure that the group wrestles with creating a moral culture for discourse.
- Second, make sure that the group's experience of discussion is constantly monitored through a classroom assessment or action research.
- Third, exercise teacher power to deconstruct and challenge structural power relations that interfere in equal discussion and equal learning.

Reference

Brookfield, S. D. 2001. A political analysis of discussion groups: Can the circle be unbroken? In *Power in practice: The struggle for knowledge and power in society,* ed. R. M. Cervero., A. L. Wilson., and Associates, 206-225. San Francisco: Jossey-Bass.

__________. 2000. The concept of critically reflective practice. In *Handbook of adult and continuing education,* ed. A. L. Wilson, and E. R. Hayes, 33-53. San Francisco: Jossey-Bass. Carli, L. L. 1990. Interpersonal relations and group processes. *Journal of Personality and Social Psychology* 59 (5): 941-951.

Cunningham, P.M. 2000. A sociology of adult education. In *Handbook of adult and continuing education,* ed. A. L. Wilson and E. R. Hayes, 573-591. San Francisco: Jossey-Bass.

Dellinger, B. 1995. *Critical discourse analysis*. Available online at http://users.utu.fi/bredelli/cda.html

Fahy, P. J. 2002. Use of Linguistic Qualifiers and Intensifiers in a Computer Conference. *The American Journal of Distance Education* 16 (1): 5-22.

Fairclough, N. 1995. *Critical discourse analysis: the critical study of language*. London and New York: Longman.

Fairclough, N., and R. Wodak. 1997. Critical discourse analysis. In *Discourse as Social Interaction,* ed. T. A. van Dijk, 258-284. London, UK: Sage.

Grob, L. M., R. A. Meyers, and R. Schuh, R. 1997. Powerful/powerless language use in group interactions: Sex differences or similarities? *Communication Quarterly* 45 (3): 282-303.

Hart, M. 2001. Transforming Boundaries of Power in the Classroom: Learning from La *Mestiza*. In *Power in practice: The struggle for knowledge and power in society,* ed. R. M. Cervero., A. L. Wilson., and Associates, 164-183. San Francisco: Jossey-Bass.

Joyce, S. 2001. The creation of bent knowledge: How lesbian, gay, and bisexual youth negotiate and reconfigure homophobic and heterosexist discourse. *Information Research* 6 (2). Available online at http//InformationR.net/ir/6-2/ws4.html

Jun, J. and J. Park. 2003. Power relations within online discussion context: Based on adult international students' perspective and their participation in the learning context. In *Proceedings of the 44th annual Adult Education Research Conference (AERC),* ed. J. M. Pettitt and R. P. Francis, 193-198. San Francisco: San Francisco State University.

Kollock, P., P. Blumstein, and P. Schwartz. 1985. Sex and power in interaction: Conventional privilege and duties. *American Sociological Review* 50: 34-46.

McAllister, C., and E. Ting. 2001. Analysis of discussion items by males and females in online college courses. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA. ERIC, ED458237.

Silverman, D. 1998. Analyzing conversation. In *Researching society and culture,* ed. C. Seale, 261-274. London: Sage.

Tisdell, E. J. 1993. Interlocking systems of power, privilege, and oppression in adult higher education classes. *Adult Education Quarterly* 43 (4): 204-226.

van Dijk, T. A. 1993. Principles of critical discourse analysis. *Discourse & Society* 4 (2): 249-283.

Williams, S. W., K. Watkins, B, Daley, B. Courtenay, M. Davis, and D. Dymock. 2001. Facilitating cross-cultural online discussion groups: Issues and challenges. In *Proceedings of the 42nd annual Adult Education Research Conference (AERC),* ed. O. Smith, J. M. Dirkx, P. L. Eddy, P. L. Farrell, and M. Polzin. East Lansing: Michigan State University. Available online at http://www.edst.educ.ubc.ca/aerc/2001/ab2001.htm

Wilson, A. L., and R. M. Cervero. 2001. Adult Education and the Struggle for Knowledge and Power: Practical Action in a Critical Tradition. In *Proceedings of the 42nd annual Adult Education Research Conference (AERC),* ed. O. Smith, J. M. Dirkx, P. L. Eddy, P. L. Farrell, and M. Polzin. East Lansing: Michigan State University. Available online at http://www.edst.educ.ubc.ca/aerc/2001/ab2001.htm

Using Computer Generated Reminders as Time Management Support to Influence Assignment Completion Rates and Course Completion in an Online Master's Program

Amy Kellogg
Barbara Lockee
Virginia Tech

Abstract

This study investigated the influence of computer-generated reminders on the rate in which distance learners submitted assignments and completed courses. The computer-generated reminders, sent via email, served as a time management support strategy. Participants were randomly divided into two groups: control and treatment. Both groups received a list of target due dates for course assignments. The control group did not receive reminders. The treatment group received reminders when they failed to submit an assignment by a target due date. The results indicated no significant difference between the control and treatment groups in terms of assignment submission rates and course completion rates. However, results of this study did reveal that the number of assignments in a distance course influences the timeliness of assignment completion and the likelihood of course completion.

Introduction

Student procrastination and attrition are common problems cited by distance instructors (Wilkinson & Sherman, 1989: 1990). As students delay the completion of assignments, they are more likely to drop out of a course or program as they fall further behind in their work. Additionally, inability to manage time is often cited by students in distance learning as a reason for not completing coursework (Bernt & Bugbee, 1993; Brown, 1996; Morgan & Tam, 1999; Saba, 2000).

Most unsuccessful distance learners are not trained in the skills necessary for studying at a distance as they have typically only participated in traditional learning environments throughout their academic careers (Morgan & Tam, 1996). Traditional, face-to-face instruction seems to naturally incorporate time management strategies which lessen the likelihood of procrastination. Generally, face-to face instruction has fixed semesters in which work must be completed, there are regular class meetings each week for a set time period, and there are also weekly assignments to complete and submit (Bernt & Bugbee, 1993). Distance courses may not have such strategies built in to their design, particularly asynchronous, self-paced learning environments.

Procrastination and course attrition are not recent obstacles to self-paced course completion. Many of the past studies related to these issues were conducted as Keller's (1968) Personalized System of Instruction courses (PSI) became increasingly popular. While the self-paced attribute of PSI courses was, and still is, attractive to many learners, such courses experience high rates of student procrastination (Born & Whelan, 1974; Bijou et al., 1976) and attrition. With the increasing popularity of asynchronous, web-based instruction, which is generally self-paced, there is a growing need for new research in the area of student pacing interventions.

Studies have been conducted in which various types of pacing contingencies were implemented within self-paced courses in order to influence the rate at which a student progresses through a course. Some pacing contingencies have been found to have a positive influence on student assignment completion rates and patterns. Such contingencies include awarding points for assignments submitted early in the week (e.g. Bijou, Morris, and Parsons, 1976), deadlines imposed uniformly throughout the term (e.g. Glick & Semb, 1974), and target deadlines for assignment submittals (e.g. Miller, Weaver, & Semb 1974). Reminders about deadlines have also been used to positively influence student submission patterns (e.g. Belawati, 1998).

The present study examined the influence of computer-generated email reminders on the rate in which distance learners submit assignments and course completion rates. Computer-generated reminders were used not only to motivate learners to move forward in their assignments, but also encourage students to monitor their own progress. By distributing computer generated reminders to students when they failed to submit an assignment by a target due date, this study investigated the effects of such reminders as a time management support strategy. There are two research questions for this study:

1. Do computer generated email reminders influence assignment submissions rates?
2. Do computer generated email reminders influence course completion rates?

Methodology

Participants and Setting

Participants of this study included learners enrolled in one of twelve courses in an online master's program Spring 2003. Both novice and experienced distance learners comprised the sample. The sample consisted of 82 learners officially admitted into the Master's program and who had registered for courses by the first day of Spring 2003 classes. If a student registered for a course after the start of the term, they were not included in the sample. In addition, any ITMA student graduating in May 2003 was not included in the sample.

As is the case with many online programs, the program examined in this study had a relatively high percentage of students who do not complete courses each semester. When students do complete courses, a large majority of them tend to procrastinate in completing and submitting their work until the end of the term. As research has shown, this end of the term rush causes a strain on faculty in terms of grading and submitting final grades (Born & Moore, 1978). Additionally, appropriate feedback cannot be given if learners submit assignments in bulk without pacing themselves. The purpose of this study is to examine the influence of computer generated assignment reminders on assignment submission rates and course completion.

Courses in this online master's program incorporate many strategies suggested in distance learning literature that help distance learners be more successful. For example, a list of suggested due dates for assignments are given to all learners (Moore & Kearsley, 1996; Saba, 2000). The use of target deadlines is a pacing strategy that does not compromise the popular attributes of self-paced learning, yet these types of deadlines have been effective in increasing student assignment completion rates (Bijou, 1976; Glick & Semb, 1978). In addition to serving as a pacing strategy, the computer-generated reminders are a strategy to help influence self-monitoring and time management among distance learners.

Participants of this study are students enrolled in at least one of twelve courses, all offered through the online master's program. The number of required assignments varies by course, from four to 13. Most courses offered through this program are project-based. Learners are able to apply concepts within each course to a project relevant to their own needs. The final products of each course are then combined into an online portfolio.

Variables

There were two dependent variables for this study; assignment completion rate (the number of days an assignment was submitted past the target due date) and course completion. There were also two independent variables in this study; the use of the email reminder and number of assignments per course. The "number of assignments" variable categorized courses based on the number of assignments required for submission. The first assignment number group, "low", included courses consisting of four to six assignments. "Medium" was the second assignment number group and included courses having seven to eight assignments. The third assignment number group was "high" and included courses requiring 13 assignments.

Treatment

All learners comprising the sample for this study were entered into a database where they were randomly assigned to two groups, A and B. Group A, the control group, did not receive computer-generated reminders when assignment due dates were not met. Group B was given the treatment of a reminder sent via email the day after a student failed to meet a suggested due date for an assignment.

Both groups were given a list of target deadlines within the course syllabus, as is the normal with program procedure. Typically, assignments are not due the first week that classes begin, due to students dropping courses and adding courses. However, by the end of the second week, the first assignment should be completed and submitted. From then on, most courses require one assignment to be submitted each week or every other week.

Final participants for each group were not identical, with group A containing 42 participants and group B containing 40 participants. Discrepancies existed due to registered students deciding not to begin a course. Students must have submitted at least one assignment in a course to be included in this study.

A database (where assignments are submitted by students) tracked the date assignments were submitted by each student. The database compared the suggested due date of each assignment with the date the assignment was submitted. Group B participants received the treatment of computer-generated reminders if they failed to submit an assignment by a suggested deadline. It was possible for a participant in the treatment group to not receive a reminder each week, or any time during the academic term if they did not miss a suggested due date.

As suggested by Moore and Kearsley (1996), the reminders were written in a personal tone to reflect the personality of graders within the program. Reminders were a form of feedback to the learners, alerting them to their progress through a course. The content of the reminders was simple and brief. The wording of the email did not change throughout the semester (see Appendix A). The technological infrastructure in place for the distance program allowed for emails to be automatically generated and emailed to learners who did not make the suggested assignment deadline. The email was typically sent 24 hours after the suggested due date.

Data Analysis

The number of assignments required for submission varied between course: from four to 13. A second independent variable was produced to address this difference during analysis. Courses were categorized based on the number of assignments: low (4-6), medium (7-8), and high (13). To determine if the number of assignments had an influence on the dependent variables of assignment completion rate and course completion, four analyses were conducted for this study:

1. Differences in assignment completion rate between control and treatment groups.
2. Differences in assignment completion rate among assignment number groups.
3. Differences in course completion between control and treatment groups.
4. Differences in course completion among assignment number groups.

Assignment completion rate analysis

Courses varied in the number of assignments required, the amount of time given to complete an assignment varied by course. To account for the discrepancy in time allowed to complete assignments, it was necessary to use a ratio to calculate the assignment completion rate. Dividing the number of days a student took to complete an assignment by the number of days allowed to complete the assignment yielded the assignment completion rate. The ratio allowed for a more accurate analysis of assignment submissions among courses. The smaller the completion ratio, the earlier the assignment was submitted. A two-way ANOVA, or two-by-two factorial design, was used to analyze assignment completion rate based on number of assignments.

Course completion analysis

Course completion was the second dependent variable examined in this study. Students were registered for one, two, three, or four courses. Completion of all required assignments for a course established course completion. For example, if there were 13 assignments required in a course, the student had to complete all 13 to complete the course. If a single student was registered for three courses, it was possible for them to complete one of three, two of three, three of three, or none of the courses. Chi-square was used examined differences in course completion between groups. In addition, a two-by-two factorial design (two-way ANOVA) examined differences in course completion between treatment groups and among assignment number groups. Conducting the analysis in this way examined the influence that the number of assignments in a course could have on course completion.

Results

The two dependent variables for this study were assignment completion rate and course completion. The two independent variables for this study were the use of an email reminder as a treatment and the number of assignments grouped into three categories, low (4-6), medium (7-8), and high (13).

There were two research questions in this study:

1. Do computer generated email reminders influence assignment completion rates?
2. Do computer generated email reminders influence course completion rates?

Assignment completion rate

The sample for this study was comprised of 82 students enrolled in from one to four of twelve courses offered through an online Master's program. Participants were assigned to one of two groups: Group A, the control group and Group B, the treatment group. The treatment group received computer generated reminders via email if they failed to meet a target deadline for submitting an assignment.

The number of assignment submissions varied by course, from four to thirteen assignments. Due to this discrepancy in assignment numbers, courses were categorized based on the number of assignments required for submission. This categorical variable was one of the two independent variables for this study. Assignment number group "low" included courses with four to six assignments. Assignment number group "medium" included courses containing seven to eight assignments. Assignment number group "high" included courses containing 13 assignments. A total of four analyses were conducted for this study:

1. Differences in assignment completion rates between control and treatment groups.
2. Differences in assignment completion rates among assignment number groups in control and treatment groups.
3. Differences in course completion between control and treatment groups.
4. Differences in course completion among assignment number groups in control and treatment groups.

To examine the assignment completion rate, a ratio was calculated by dividing the number of days a student took to complete an assignment by the number of days allowed to complete the assignment. The resulting number was used to represent the participants' assignment completion rates.

The first research question examined the influence of email reminders on assignment completion rates. A one-way ANOVA was used to analyze the data. This study hypothesized that the treatment group would submit assignments closer to the targeted due dates than the control group. While this study did result in the treatment group submitting assignments closer to the targeted due date as compared to the control group, the difference was not significant, $F(1)=.047$, $p>.05$ (see Table 3). Table 4 presents the table of means for assignment completion rate and treatment group.

Table 3 *Analysis of Variance for Assignment Completion Rate Between Groups*

	df	F value	P value
Assignment completion rate	1	.047	.828

significant p < .05

Table 4 *Means and Standard Deviation for Assignment Completion Rate Between Groups*

Group	N	Mean	SD
Control	450	1.2604	3.26520
Treatment	479	1.2184	2.60899

*significant p < .05

However, a two-way ANOVA revealed a significant difference, $F(1,2) = 19.412$, $p < .05$, in assignment completion rate between assignment number groups (see Table 5). Significance existed among all three assignment groups. Means for assignment completion rates revealed that the fewer assignments within a course the later assignments were submitted (see Table 6). The more assignments a course contained, the closer assignments were submitted to the suggested deadline. This was true for both the control and treatment groups.

Table 5 *ANOVA Table: Treatment and Assignment Number*

	df	F value	P value
Treatment	1	.041	.840
Assignment number group	2	19.412	.000*
Treatment **x** Assignment number group	2	3.031	.0503

* significant p < .05

Table 6 *Means and Standard Deviation for Assignment Completion Rate and Number of Assignments*

Group	Assignment number	Mean	Std. Deviation	N
Control	Low (4-6)	2.4978	3.06027	92
	Medium (7-8)	1.4922	2.13916	192
	High (13)	.3066	4.08883	166
	Total	1.2604	3.26520	450
Treatment	Low (4-6)	2.1590	1.93773	61
	Medium (7-8)	1.2659	3.04936	126
	High (13)	1.0014	2.48573	292
	Total	1.2184	2.60899	479

Course completion

The second research question examined the influence of email reminders on course completion. Chi-Square was used to analyze course completion by treatment group. Analysis revealed no significant difference between groups (p = .124). The email reminder did not have a significant influence on course completion (see Table 7). Crosstabs are presented in Table 8.

Table 7 *Chi-Square Summary for Dependent Variable: Course Completion and Independent Variable: Treatment*

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.795(b)	1	.180		
Fisher's Exact Test				.195	.124
N of Valid Cases	140				

*significant p < .05

Table 8 *Crosstab Table for Course Completion in Control and Treatment Groups*

Group			Course completion		Total
			Not complete	Completed	
	Control	Count	25	50	75
		Expected Count	21.4	53.6	75.0
		Std. Residual	.8	-.5	
	Treatment	Count	15	50	65
		Expected Count	18.6	46.4	65.0
		Std. Residual	-.8	.5	
Total		Count	40	100	140
		Expected Count	40.0	100.0	140.0

When the course completion was analyzed by assignment number group, two-way ANOVA revealed a significant difference, $F(1,2) = 3.225$, $p < .05$ (See Table 9). All three of the assignment number groups (low 4-6, medium 7-8, high 13) were significantly different from one another in terms of course completion. In other words, there was significance between the low and medium assignment groups, the low and high assignments groups, and the medium and high assignment groups. While the treatment did not have an impact on course completion, the number of assignments within a course did impact course completion.

Also, examining course completion means among assignment groups revealed medium assignment number (7-8 assignments) courses had a higher course completion rate than did low (4-6 assignments) or high (13assignments) assignment number courses. This was true for both control and treatment groups. These means are presented in Table 10.

Table 9 *Two Way Analysis of Variance for Course Completion and Number of Assignments*

Source	df	F	Sig.
Treatment Group	1	1.910	.169
Assignment Number Group	2	3.225	.043*
Treatment Group **x** Assignment Number Group	2	1.020	.363

*significant p < .05

Table 10 *Table of Means and Standard Deviation for Course Completion and Number of Assignments*

Group	Assignment Number Group	Mean	Std. Deviation	N
Control	Low (4-6)	.6842	.47757	19
	Medium (7-8)	.7742	.42502	31
	High (13)	.5000	.51075	24
	Total	.6622	.47620	74
Treatment	Low (4-6)	.6429	.49725	14
	Medium (7-8)	.9000	.30779	20
	High (13)	.7419	.44480	31
	Total	.7692	.42460	65
Total	Low (4-6)	.6667	.47871	33
	Medium (7-8)	.8235	.38501	51
	High (13)	.6364	.48548	55
	Total	.7122	.45436	139

Discussion

Pacing studies of the 1960's and 1970's typically found a steadier rate of assignment completion when pacing contingencies were implemented throughout a semester (Lazar et al., 1977; Miller et al., 1974). Although the treatment group did submit assignments closer to the targeted due date as compared to the control group, the difference was not significant. The treatment employed in this study did not influence assignment completion rates. Although past studies did find a steadier rate of assignment completion when a treatment was implemented, course completion rates between groups was not found to be significantly different (e.g. Glick & Semb, 1978). The same is true for this study. The treatment did not significantly influence course completion.

With experience in the ITMA program comes the knowledge that suggested deadlines are in fact suggested deadlines. Learners new to the ITMA program may panic when they miss a deadline, assuming these deadlines are fixed as is generally the case in face-to-face environments. Learners experienced in the ITMA program are typically aware of the policies and understand that as long as they submit all work prior to the end of the term deadline, they will be awarded a final grade. The randomization used in this study ensured heterogeneous control and treatment groups so these differences were not observable. Each group had a fairly even blend of experienced and novice distance learners.

Serban (2000) conducted research that found policies such as number of distance courses learners are allowed to enroll in influence student success. As is true with traditional, on-campus courses, the more online courses a student enrolls in during a semester, the more likely that student will withdraw.

Unexpected Findings

Although this study did not examine the influence of number of courses on assignment completion, it did reveal a significant difference among assignment number groups (low, medium, and high) for both assignment completion rate and course completion. This study revealed that the more assignments in a course, the closer assignments were submitted to the suggested deadline. Students submitted assignments more days past the suggested due dates when fewer assignments were in a course. This research is evidence that the number of assignments required in a distance course influences the timeliness of assignment completion and the likelihood of course completion.

Implications

In this study a significant difference in assignment completion rate and course completion was found between assignment groups. The findings of this study are important to distance course design. Due to the problems of procrastination and attrition in distance learning, distance education would benefit from more research on how the number of assignments in a course influences the assignment completion rate and course completion. This study found that the more assignments in course, the m

Research has shown that procrastination, and potentially attrition, can be reduced by implementing such strategies as providing orientation to distance learners, contacting students via email who are not submitting assignments, and increasing frequency and immediacy of feedback (Moore & Kearsley, 1996; Coggins, 1989; Kirkwood, 1989; Belawati, 1998; Eisenberg & Dowsett, 1990; Wilkinson & Sherman, 1990). These issues are not within the control of the student, but the institution and could potentially be addressed with minor policy and process changes to grading procedures. Student support cannot be overlooked in a distance program. This type of support may require the training of personnel, as many institutional staff may not be knowledgeable about providing this type of support to distance learners.

Learners new to the distance environment may be more appreciative of the reminders even though the reminders were computer generated. Procedures of this sort would require a more in-depth analysis of learners prior to, or at the beginning of each term. Such methods could then be removed or replaced with new strategies of motivation and approaches that foster self-monitoring as students become more experienced in a program.

References

Belawati, T., (1998). Increasing student persistence in Indonesian post-secondary distance education. Distance Education, 19 (1), 81-108.

Bernt, F. & Bugbee, A. (1993). Study practices and attitudes related to academic success in a distance learning programme. Distance Education, 14 (1), 97-112.

Bijou, S., Morris, E., & Parsons, J. (1976). A PSI course in child development with a procedure for reducing student procrastination. Journal of Personalized Instruction, 1(1), 36-40.

Born, D. & Whelan, D. (1974). Some descriptive characteristics of student performance in PSI and lecture courses. In J.G. Sherman (Ed.), PSI: Personalized System of Instruction 41 Germinal Papers (153-159). London:
W.A. Benjamin, Inc.

Coggins, C. C. (1989). Preferred learning styles and their impact on completion of external degree programs. In M. G. Moore & G. C. Clark (Eds.), Readings in distance learning and instruction (pp 1-12). The Pennsylvania State University, USA.

Eisenberg, E. & Dowsett, T. (1990) Student drop out from a distance education project course: A new method of analysis. Distance Education, 11 (2), 231-253.

Glick, D., & Semb, G., (1978). Instructor set pacing contingencies verses the absence of such contingencies in a personalized university course. Journal of Personalize Instruction, 3(3), 131-138.

Lazar, R., Soares, C., & Terman, M. (1977). The interquiz interval: A measure of student self-pacing.

Journal of Personalized Instruction, 2 (1), 25-27.
Lloyd, K. & Knutzen, N. (1969, Summer). A self-paced programmed undergraduate course in the experimental analysis of behavior. Journal of Applied BehaviorAnalysis, 2, 125-133.
Miller, L., Weaver, F., & Semb, G. (1974). A procedure for maintaining student progress in a personalizes
university course. Journal of Applied Behavior Analysis, 7 (1), 87-91.
Moore, M. G., & Kearsley, G. (1996). Distance education a systems view. Belmont, CA:Wadsworth Publishing Company.
Morgan, C. & Tam, M. (1999). Unraveling the complexities of distance education student attrition. Distance Education, 20 (1), 96-108.
Saba, F., (1999). Designing instruction for the distant learner. Distance Education Report, 3(4), 2-7.
Semb, G., Glick, D., & Spencer, R. (1979). Student withdraws and delayed work patterns in self-paced psychology courses. Teaching of Psychology, 6 (1), 23-25.
Serban, A. M. (2000). Evaluation of fall 1999 online courses. (ERIC Document Reproduction Services No,
ED453868).
Wilkinson, T. & Sherman, T. (1990, December). Distance education and student procrastination. Educational Technology, 24-27.
Wilkinson, T. & Sherman, T. (1989). Perceptions and actions of distance educators on academic procrastination. The American Journal of Distance Education, 4 (3), 47-55.

Appendix A

Email Reminder

Dear [student's first name],
The target submission date for [assignment number and title] was [suggested due date]. While there is no penalty for submitting assignments past the suggested due date, we want to help you stay on track to finish the course. If you have already submitted your assignment, please disregard this reminder. If you have any questions or problems with the assignment, please feel free to contact itmahelp@vt.edu.
-Amy

Understanding Characteristics of Lurkers in E-learning

Eunok A. Kim
University of Illinois at Urbana-Champaign

Abstract

The purpose of this study was to explore what factors are different among the following groups with different participation patterns: Active participants, Lurkers, and Inactive participants. Throughout the literature review, the following factors were suggested to be related to learner participation and examined as different factors among participation groups: computer skills, locus of control, study flow, class flow, computer flow, physical environment, and support.

The sample for this research was an e-learning course at the university level. Participation in e-learning courses for this study was measured by the number of pages accessed by students, log-on time, and the number of message postings. Log-file, questionnaire, and observation data of the class were collected for analysis. The data were analyzed with multiple regression method using orthogonal coding.

There was no difference in seven factors between the active participants and the lurkers. However, there were some differences in contrast among the active participants and the lurkers VS the inactive participants in class flow, support and physical environment. Understanding characteristics and differences of lurkers from active participations would contribute to understanding e-learners.

Introduction

E-learning based on network technology is known to have potentials that would allow for learners to be more self-motivated and active participants in their own learning since it allows learners to have more autonomy in setting the pace of their own learning and higher interactivity in learning regardless of time and space than in face-to-face environment. While these possibilities and potentials have been addressed in literature, researches have found influences of e-learning are multi directed and depend on various factors such as learners, instructional styles, and contextual factors.

Learner participation is a crucial issue in e-learning. It is a common proposition that e-learning cannot be successful without learners' "active" participation. Te proposition need to be examined with further probing questions: How can we define participation in e-learning setting? What kinds of participation exist in e-learning? Are they different from those in face-to-face environment? What participation is needed for e-learning to succeed? The research started with these questions and the researcher attempted to examine what factors are different among learner groups with different participation patterns and to explore what the implications of the results are.

Mason (1994) addressed the 'thirds theory' which suggests that students fall into three distinct group: those who actively participate, those who read messages but do not post messages and those who take no part. Similarly, Taylor (2002) categorized participants into the following three sub-groups that differ in terms of the participation patterns in accessing, and contributing to a discussion group: Proactive participation group (workers), Peripheral participation group (lurkers), and Parsimonious participation group (shirkers). Lurkers are known as "people who post occasionally or not at all but are known to read the groups' postings regularly." (Online Jargon Dictionary, available at http://www.netmeg.net/jargon) Lurking is also known as one of the most popular participation patterns in the online community. For example, according to online demographic researches, lurkers varies size ranging from 9% to 99% in discussion groups and reported to make up 90% of several virtual groups (Nonnecke & Preece, 2000; Nonnecke & Preece 2001; Mason,1999).

This research identified the following three groups based on combined measures of the number of pages accessed by students, log-on time, and the number of message postings: the active participants, the lurkers, and the inactive participants.

The purpose of this study was to explore what factors are different among the active participants, the lurkers, and the inactive learners in e-learning. It was examined whether or not factors, including locus of control, computer skills, study flow, class flow, computer flow, physical environment, and support, were different among groups with different learner participation patterns. The researcher wanted to know what are different among participants with different participation patterns. Through exploring differences among participation groups, this study would contribute to the understanding of learners and their participation in e-learning environment.

Research Question

What factors are different among participant groups (the active participants, the lurkers, and the inactive participants)?

Limitations

This study has the following limitations:

1. The generalizability of the results is limited to populations that are similar to the sample studied. The course studied is a web-based class supplemented by face-to-face instruction at the post-secondary level. More research needs to be done on courses that are only delivered through online and on classes of various school levels.

2. There can be other criteria for identifying participation patterns besides log-on time, the number of pages accessed, and the number of messages posted.

3. The criteria used to divide the three groups were set up according to the researcher's own judgment. The criteria could vary.

Definitions of Terms

- Learner participation in e-learning: Degree of learner participation in e-learning was operationally defined the number of pages accessed, log-on time, and the number of message postings in this research.
- Flow: "Flow is a dynamic state- the holistic sensation that people feel when they act with total involvement" (Csiksszentmihalyi, 2000). It is not a desire to achieve external goals but an auto-telic experience, in which the activity itself becomes the goal. It is characterized as a playful and exploratory experience. Three types of flow, including study flow, class flow, and computer flow were examined in this study.
- Lurkers: Lurkers in computer-mediated communication are people who post occasionally or not at all but are known to read the groups' postings regularly

Theoretical Framework

This review of related literature first focused on factors related to learner participation in e-learning. Then, the literature on lurkers in computer-mediated communication(CMC) was briefly reviewed.

Factors Related to Learner Participation in E-Learning Courses

While there is a great deal of research in e-learning generally, there are few studies of learner participation specially. For that reason, factors influencing learner participation were drawn from studies in the related field of computer-mediated communication.

Participation in e-learning courses might be influenced by a number of factors. In this research, the following seven factors are suggested and examined with relation to learner participation in e-learning through the literature review; locus of control, computer skill, study flow, class flow, computer flow, physical environment, and support. These were selected because literature highly suggested them to be factors that influence the use of computer-mediated communication.

While face-to-face communication is considered 'natural' communication, computer-mediated communication requires basic knowledge and skills of computers and networking(Kim, 1996). Computer skills refer the ability to perform application programs such as word-processor, multi-media programs, and programs used for the internet. Much literature shows that computer skills influence the frequency of usage and the pattern of computer mediated communication (Grabowski et al, 1990; Jung & Choi, 1998; Kayany & Rowley, 1994; Kim, 1996; Lim, 1996).

Locus of control is "a personality construct referring to an individual's perception of the locus of events as determined internally by his/her own behavior vs. fate, luck, or external circumstances" (Choi, 1996; Rotter, 1996). If one believes that he or she is able to control one's life and that his or her circumstances depend on their own decisions, it is said that one's locus of control is internal. On the contrary, if one believes situations are dependent on things outside of their control, one's locus of control is said to be external.

Locus of control is known as an influential factor in learning and using new computer programs, and in anticipating computer mediated communication (Bruning, 1992; Howard, 1986). Learners with an internal locus of control tend to have more possibilities and opportunities to improve performance and have more confidence in computer-mediated communication.

Flow is a special kind of experience defined as a "dynamic state that people feel when they act with total involvement" (Csikszentmihalyi, 1990). The flow theory was developed by a psychologist Mihaly Csikszentmihalyi in 1970s and has been applied in various areas such as education, communication, and human-computer interactions (Carli, 1986; Csikszentmihalyi, 1990).

Flow is a figurative term, which was frequently described in interviews with people who were doing what they really liked. Flow can be experienced in diverse activities. One of the important implications of flow experience is that this experience can be experienced in any daily life activity ranging from leisure and creative projects to works (Csikszentmihalyi, 1975).

There are common structural characteristics of flow experiences. One of the characteristics of flow experiences is that it is pursued not by the desire for external goals but by the activity itself.

Flow comes about as the result of the balance between the challenge presented and the skills an individual has. If the skills an individual possesses are too limited to meet the challenge, this elicits anxiety. On the contrary, if the activity is too easy, then boredom arises. In other words, low level challenges cause boredom and challenges of extremely high level beget anxiety. As skills improve, the interaction between skills and challenges presented lead one to seek challenges of even greater difficulties (Csikszentmihalyi & Csikszentmihalyi 1993). Therefore all kinds of activities have the possibility to develop flow experience; however, unless challenges or skills are made more difficult, no activity can maintain a flow experience. The task needs to spiral in complexity.

Flow makes people take on challenges and develop their abilities (Clarke & Haworth, 1994; Csikszentmihalyi, 1990; Miller, 1974). Increasing levels of challenges are related to internal rewards (Csikszentmihalyi, 1988). One perceives the activity not just as a tool for another purpose but as the goal itself which is joyful and meaningful.

Flow facilitates learning since one practices and develops one's skill through exploratory activities that elicit a sense of flow (Miller, 1973). Flow is also a useful concept in explaining the interaction between humans and computers (Csikszentmihalyi, 1990; Ghani, 1991; Malone, 1981; Turkle, 1984, Trevino & Webster, 1992).

For this study, three kinds of flow were examined, the first of which is study flow, and is defined as the flow experience in studying on one's own (Harju & Eppler, 1997). Because e-learning is mostly pursued by individual efforts, study flow has been suggested as a related factor to learner participation.

Secondly, class flow, which can be described as the content, and the subject matter of the class, was examined in this study. Because instructions are delivered through computer networks in e-learning, computer flow was also investigated.

Finally, computer flow was also included. Computer flow is defined as the flow experience while working on computers, and which could mean any thing from checking email to playing games.

Literature shows that physical environment is related to the amount of computer-mediated usage (Markus, 1987; Grabowski et al, 1990; Kaye, 1987, Steinfield, 1986). Physical environment means the infrastructure to access web-based instruction such as available computers, network, speed and quality of connections

Finally, psychological support and help were suggested to be factors that influence the participation of online communities and the usage of computer-mediated communication (Hiltz, 1984; Kayany &Rowley, 1994; Riel & Levin, 1990). Support was measured by whether or not there were people available such as instructors, teaching assistants, and colleagues for technical help and guidance.

Lurkers in E-Learning

Lurkers in computer-mediated communication refer to people who frequently visit certain online sites, spend considerable time there, but seldom post their own messages. Lurking is a particular participation pattern that utilizes the fact that participants cannot see one another, which is one of characteristics of online communities (Romiszowski & De Hass, 1989).

Lurkers in e-learning are different from people who seldom participate in the e-learning system. Lurkers often use e-learning systems and spend a considerable amount of time there. Lurkers are different from active participants only in that they seldom post their messages. Lurkers are not easily recognized by observers or researchers because of the virtual characteristics of online community.

Choi (1996) suggested that people with external locus of control tend to be lurkers. These people primarily use computer-mediated communication to receive information. Poor typing and writing skills are known to be related to lurkers (Schmitz & Fulk, 1991).

Research Methods

For this study, a quantitative research framework was adopted. The learner participation in the e-learning course was measured by the number of pages accessed, log-on time, and the number of messages posted. Other factors such as locus of control and environment were measured by a questionnaire.

The sample course studied was selected from one of online classes at Seoul National University (SNU), Korea. The sample class also included traditional face-to-face instruction as a small part of the course. The sample class was a sociology class which was open to all SNU students regardless of their majors and programs of study at SNU. It was a web-based class that used TopClass as the class platform, and it was supplemented by face-to-face instruction when the guest speakers were invited to lecture. This class only gave two grades; S(satisfactory) or U(unsatisfactory). Of the 232 total students who registered in the sample class, 139 responded to the questionnaire. Data were collected for approximately 6.5 weeks. The number of pages accessed and student log-on time were collected from the log files in the server computer. The number of message postings was collected through observation of web boards in the online class. Questionnaires were distributed and collected in the face-to-face class site. 139 responses were analyzed.

The number of message postings was used to divide the lurker group from the active participant group. It was tracked through observation of the e-learning course and identified with users. The instructor assigned three messages to post online. The number of messages posted by the instructor, and learners totaled 1039. The mean of posted messages was 5.00 and the standard deviation was 4.10.

The survey questionnaire was made up of questions related to factors including computer skills, locus of control, flow (study, class, and computer), physical environment, and support. The variable, scale, and the number of items comprising each scale are presented in table 3-1.

Table 1. *Description of Variables and Scales Used in the Study*

Variable Name	Scale/ Source of Items	Number of Items
Computer Skill	Kim(1998)	8
Locus of Control	Levenson(1974), Choi(1996)	12
Study flow	Harju & Eppler(1997)	8
Class Flow	Harju & Eppler(1997)	8
Computer Flow	Researcher developed	8
Physical Environment	Jung & Choi(1998)	3
Support	Jung & Choi(1998)	3

Study and class flow items from Harju and Eppler's (1997) questionnaire were adapted from Csikszentmihalyi's (1988) flow questions. Study flow variable consist of eight items including statement such as "I clearly know what I am supposed to do while studying or learning on my own", and "I would study and learn on my own even if I didn't have to do it".

The researcher developed computer flow items using equivalent format used for the study flow and class flow items. For instance, the following items were used measuring computer flow. "I got involved in doing something using the computer", "I feel I can handle the demands of the situation while doing something using the computer".

Data were analyzed with the statistical methods of multiple regression using a SPSS Win v.10.0 package.

Research Findings

While the number of pages accessed and log-on time could be thought as an indicator of class attendance and participation through listening, the message-posting can be thought of as vocal participation in e-learning.

Participant groups were named as active participants, lurkers, and inactive participants according to participation patterns and divided using the following steps. First, by definition, the active participants and lurkers need to have a high number of pages accessed and a long log-on time. The researcher selected a group of people who are above medians in both number of accesses and log-on time based on medians.
Then active participants and lurkers were divided by the median of message posting from the group. Inactive participant were identified as people who rank below medians in all of the number of access, log-on time, and message posting in this study.

Total Samples

▼

Above median of number of pages accessed &
Above median of log-on time

▼

Message posting above median **(Active Participants)**	Message posting Below median **(Lurkers)**	below median in the number of access, log-on time, and message posting **(Inactive Participants)**	

Figure 1. Definition of participant groups

The sample sizes of each group are as follows. Researcher planned two comparisons. The first comparison examined there was any difference between the active participants and the lurker groups (Comparison 1). The second comparison is about group difference between the active participants and, the lurkers VS the inactive participants (Comparison 2).

Table 2. *Sample Size and Orthogonal Comparison*

	Group 1 Active Participants	Group 2 Lurkers	Group 3 Inactive participants
SampleSize	35	23	69
O1 Contrast	1	-1	0
O2 Contrast	½	½	-1

Since each group has unequal sample size, orthogonal coefficients were set as follows:

Table 3. *Orthogonal Comparison Coefficients*

	Group 1 Active Participants	Group 2 Lurkers	Group 3 Inactive participants
Sample Size	35	23	69
O1 Coefficient	-23	35	0
O2 Coefficient	69	69	-58

The results are as follows. were in the following table.

Table 4. *Comparisons of Factors among the Three Groups*

	Contrast	B	BETA	T	Sig.
study flow	O1	9.453E-04	.032	.355	.723
	O2	1.323E-03	.146	1.640	.104
class flow	O1	3.280E-04	.010	.110	.913
	O2	2.905E-03	.279	3.219	.002**
Computer flow	O1	-3.069E-04	-.077	-.077	.939
	O2	2.056E-03	.147	1.652	.101
Computer skills	O1	-3.067E-04	-.006	-.067	.947
	O2	2.495E-03	.160	1.789	.076
locus of control	O1	2.739E-03	.097	1.075	.285
	O2	-3.541E-06	.000	-.005	.996
Support	O1	9.901E-04	.032	.358	.721
	O2	1.9663-03	.212	2.389	.018*
Physical environment	O1	-3.381E-04	-.007	-.084	.933
	O2	2.718E-03	.199	2.243	.027*

There was no difference in the seven factors (study flow, class flow, computer flow, computer skills, locus of control, support, and physical environment) between the active and the lurker participation groups.

However, there were some differences in contrast among the active participants and the lurkers VS the inactive participants in the factors of class flow, support and physical environment. The class flow, support, and physical environment in the inactive group were significantly lower than those in the active participants/lurkers. Generally speaking, while the inactive participants did not show intrinsic motivation or interest in the content of class, both the active participants and the lurkers are more interested in the subject areas of class and feel sense of control that their skills can manage the challenges presented in the subject matter. Inactive participants also reported that they did not have people to help with their e-learning class and enough equipments or networks.

It seems that the lurking learners share much more common things with the active participant group than with the group of inactive participants. Even though the lurkers could be easily misrecognized as inactive participants without special efforts such as log-file analysis, there seems to be some important difference between the lurkers and the inactive participants.

Conclusions

Lurking learners are often perceived as inactive participants by instructors since they do not post messages often and there is no 'visible' evidence that they participate in online classes. However, we need to distinguish lurkers from inactive participants. Much literature has shown that lurking is a popular participation pattern in online communities and it is different from inactive participation. The result of this empirical study also suggests that lurkers seem to be more similar to active participants than to inactive participants in the several factors that affect participation in e-learning.

There are still many questions needed to be addressed about lurking and lurkers. Why learners lurk? Why lurkers invest considerable amount of time but do not post their own messages? Do lurkers learn differently from other learners? We need to pay more attentions on lurkers and meaning of lurking. Understanding of characteristics lurkers would contribute to understanding e-learners.

E-learning is gaining in popularity. Currently, twenty five percent of companies in the U.S. use computer-based games and simulations and twenty percent use distance learning (Noe, 2002). However, while e-learning is popular in practice, few empirical studies on learners and their behaviors have been conducted. We need to be more mindful of learners and their participation patterns that make e-learning to be successful and learners to learn.

References

Bruning, S.(1992). *An examination of the social, psychological, and communication variables that influence user perceptions of computer-mediated technologies.* Unpublished doctoral dissertation, Kent State University.

Choi, S.H. (1996). *Factors related to the collaborative uses of computer-mediated communication in a graduate community: a study of electronic mail*. Unpublished doctoral dissertation, Michigan State University.

Csikszentmihalyi, M.(1975). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass.

Csikszentimihalyi, M. (1990). *Flow: The psychology if optimal experience*. New York: Harper & Row.

Csikszentmihalyi & Csikszentmihalyi (Eds). (1988). *Optimal experience :psychological studies of flow in consciousness*. Cambridge University Press.

Harju, B. L. & Eppler, M.A.(1997). Achievement Motivation, flow and irrational beliefs in traditional and nontraditional college students. *Journal of Instructional Psychology*, 2, 147-157.

Hatch, M.L. & Hayward, M.S. (1996). Computers and the college classroom: two studies of computer training and use patterns. *Proceedings of ED-TELECOM 96*. Boston. Association for the Advancement of Computing in Education.

Hiltz, S. R.(1984). *Online communities: A case study of the office of the future*. Norwood, NJ: Ablex.

Howard, G.S.(1986). *Computer anxiety and the use of microcomputers in management*. Ann Arbor, MI: UMI.

Ghani, J.A. (1991). Flow in human-computer interactions: test of a model . in J.Carey(Ed), *Human factors in management information systems: An organizational perspective*. Norwood. NJ: Ablex. 291-311.

Grabowski, B., Suciati, & Pusch, W.(1990). Social and intellectual value of computer-mediated communications in a graduate community. *ETTI,* 27, 276-283.

Jung, I.S & Choi, S.H. (1998). *Factors related to effectiveness in online training*. . Seoul :Korea Open University.

Kayany, J.M. & Rowley, M.W.(1994). Synchronous and asynchronous forum in cyberspace for theoretical dialectics. Paper presented at the AEGMC convention, Communication theory and methodology division.

Kaye, A. (1987). Introducing computer-mediated communication into a distance education system. *Canadian Journal of Education Communication.* 16(2), 153-166.

Kim, M.(1998). *Strategies to improve interactivity in hyper-text learning system* .Unpublished doctoral dissertation, Seoul : Seoul National University.

Kim, E.A. (2003). *Factors related to learner participation in e-learning.* Paper presented at the Academy of Human Resource Development 2003 Annual Conference Proceedings, Minneapolis, MN

Kim, Y.J. (1996). Cyberspace and human communication: computer-mediated communication as interpersonal communication. *Press, Society and Culture*. Korea: Younsei University.

Levenson, H.(1974). Activism and powerful others: Distinctions within the concept of internal-external control. *Journal of Personality Assessment*, 38.

Lim, H.K. (1996). *Formation and characteristics of virtual communication*. Unpublished master thesis. Seoul: Seoul National University.

Mason, R. (1994). *Using communication media in open and flexible learning.* London: Kogan Page.

Malone, T.W. (1981). Toward a theory of intrinsically motivated instruction. *Cognitive Science*, 4.

Markus, M.L.(1987). Toward a critical mass theory of interactive media-universal access,

interdependence and diffusion. *Communication research.,* 14(5), 491-511.
Noe, R.A .(2002). *Employee training & Development*, New York: McGraw Hill.
Nonnecke, B. & Preece, J.(2000). Lurker demographics: counting the silent. CHI 2000, The Hague, 2000a, ACM Press. pp.73-80.
Nonnecke, B. & Preece, J.(2001). Why lurkers lurk, Americas Conference on Information Systems.
Pedhazur,E.J. (1997). Multiple regression in behavioral research. Orlando, FL: Harcourt Brace.
Romiszowski, A.J. & De Hass, J.A.(1989). Computer-mediated communication for instruction: Using e-mail as a seminar. *Educational Technology* 29,7-14.
Riel & Levin. (1990). Building electronic communities: Success and failure in computer networking. *Instructional Science,* 19, 149-169.
Rotter,J.B.(1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*. 80(1).
Schmitz & Fulk (1991). Organizational Colleagues, Media Richness, and electronic Mail. *Communication Research.* 18(4), 487-523.
Steinfield, C.W.(1986). Computer-mediated communication in an organizational setting: Explaining task-related and socioemotional uses. In M. L. Mclauhlin(Ed), *Communication yearbook.* Vol. 9, 777-804. Beverly Hills. CA: Sage.
Taylor, J.C. (2002). Teaching and learning online: the workers, the lurkers and the shirkers. Retrieved March, 2003, from http://www.ouhk.edu.hk/CRIDAL/cridala2002/speeches/taylor.pdf
Trevino, L.K. & Webster, J. (1992). Flow in computer mediated communication: electronic mail and voice mail evaluation and impacts. *Communication Research,*19(5), 539-573.
Turkle, S.(1984). *The second self.* New York: Simon & Schuster.

Comic Strips in Instruction for Sixth Grade Science: A Case Study in Instructional Design and Development

Myonghee Kim
Nari Kim
Ray Martinez
Indiana University

Abstract

This paper discusses a systematic approach to instructional design following the ADDIE (analysis, design, development, implementation, and evaluation) model. The paper illustrates a case study in which a graduate school instructional design team in a first course in instructional design and development went through a series of steps for creating a three lesson self-study module for fifth and sixth graders to learn about cells and their parts. Particularly, this paper discusses the context and background of the design project; provides the theoretical frameworks employed by the designers; overviews the analysis, design and development, and evaluation of the instruction; and concludes with a brief synopsis of lessons learned and ideas for improvements. Focus is on illuminating what effect following the ADDIE model had on the developmental process of the instructional material, how closely each step involved in ADDIE was linked to the other steps, and what factors were taken into account in the steps. The instructional product utilizes the context of comic strips as a primary instructional vehicle, which the analysis process identified as serving to motivate and maintain learners' interest in learning. Examples of the final instructional product including student material, teacher guide, and a website are provided with explanations. This paper will be of interest to instructional designers and material developers.

Introduction

This paper discusses the instructional design process used by a team of graduate students in the instructional systems technology program at Indiana University. The design team members were students in the course "Instructional Design and Development." The course is the first instructional design course in the graduate program. It is one of the courses required of all first year students.

Two main projects were completed during the course: one on a procedure or procedural skill learning, and one on a concept learning. This paper is based on the second concept learning project.

The design team of three chose the topic of "Cells and their Parts" for the project from a predetermined list of approximately 15 choices and created a concept lesson to help fifth and sixth graders to understand the characteristics of cells, identify the main features of cell parts, and discriminate the similarities and differences between plant and animal cells.

In this paper, we aim to describe the systematic process of the analysis, design and development, and evaluation of the instruction; as well as our design decisions and accompanying rationales. We relate the effect of following the ADDIE model on each step, the systematic nature of the model, and the factors taken into account in the steps.

Theoretical Framework

The theoretical framework for designing our instructional material derived from three sources: the ADDIE model, Reigeluth's principles for teaching concept classification, and the use of comics for instructional purposes.

ADDIE Model

In the field of instructional technology, a systems approach to instructional design is valued and widely adopted. The systems approach considers all the steps involved in the instructional design as one interconnected, integrated process. Major phases of the instructional design process are analysis, design, development, implementation, and evaluation (ADDIE), which work together to bring about desired learning outcomes (Dick, Carey, & Carey, 2001).

The analysis step involves four aspects of analysis: needs analysis, learner analysis, content analysis, and context analysis. Analysis aims at determining the need for instruction in certain skills or knowledge, and also examines learner characteristics including learners' current level of skills and knowledge, preferences, and attitudes. In addition, the content to teach and the context for instruction are determined. Based on the results of analysis, specific instructional objectives are established, identifying the skills and knowledge to be learned, the conditions under which the skills and knowledge are performed, and the criteria for judging successful performance (Mager, 1997).

The design step focuses on identifying instructional strategies, instructional media, and message design plan. Particularly, instructional strategies are selected considering learner characteristics and the content to teach. This design step serves as a blueprint for the next development step.

Following the design step is the development stage in which actual instructional materials are produced and prepared. Examples of instructional materials include learner materials, teacher guide, overhead transparencies, websites, and videotapes.

The implementation step occurs when instruction is implemented using the instructional materials developed. During the implementation stage, formative evaluation is conducted to collect data to determine whether the instruction is effective and to identify ways of improving the instruction. Data from the formative evaluation are used to revise the instruction and instructional materials. After the instruction is revised according to the results of the formative evaluation, main implementation of instruction takes place. When the instruction is complete, a summative evaluation is conducted to determine the overall effectiveness of the instruction. Whereas formative evaluation is concerned with diagnosis of the planned instruction and instructional materials prior to actual implementation, summative evaluation is concerned with evaluating the value or worth of the instruction as a whole.

The five steps are closely interconnected with one another in such a way that the output of one step serves as input of the next step. Each step can be revisited and revised based on the output from a step.

Reigeluth's Principles for Teaching Concept Classification

As an instructional theory to design our instruction, we adopted Reigeluth's principles for teaching concept classification (1999). Reigeluth's principles for teaching concepts consist of several routine and enrichment tactics including presentation (of both a generality and examples), practice, and feedback. To help the learner to classify concepts in presentation, a generality should be presented which includes criterial definitions of concepts based on their critical characteristics, as well as the functional definitions of the concepts on the basis of their functions. Focusing the learner's attention and using variety in presentation can enrich a generality. With a generality, examples should be presented simultaneously for the facilitation of concept acquisition. Divergent examples of concepts and their matched non-examples should be provided in an easy-to-difficult sequence. After being provided with a generality and examples, practice should offer new divergent instances for learners to generalize the classification skills to new examples. Increasing the number of practice items and presenting them in an easy-to-difficult sequence can enrich the practice tactic. Following practice, feedback should be provided with proper guidelines based on the result of the learner's choice. A variety of representation and praise can be used as feedback for correct responses, while hints and encouragement can be provided as feedback for incorrect responses.

Use of Comics

The choice of comics as a mode of delivery for the concept lesson was based initially on the simple observation that "kids like to read comics." As motivation has been shown to be a factor in effective learning (Keller, 1983), the design team decided to investigate this form of delivery. In addition to the observation of children, two of the design team members, both from South Korea recalled from their own grade school years, that many students read instructional comics outside of classroom.

A brief review of the literature found that the use of graphics has been well discussed in instruction (Houghton & Willows, 1987; Misanchuk, Boling, & Schwier, 1999). In particular, message design has received a large amount of attention (Fleming & Levie, 1993). The design team, however, found few research articles focusing specifically on the instructional use of comic strips, despite instructional comics appearing in the United States as early as 1951 (Andree, 2001). Several informal, but practical resources were nonetheless located which alluded to the use of comics for instruction (Eisner, 1994; McCloud, 1994; Eisner, 1996; Carrier, 2001).

Through informal discussions with learners and science teachers about initial conceptions regarding comics, the team felt comfortable proceeding based on this limited literature review.

Instructional Design Process

For the purposes of this project, we completed the analysis, design, and development steps in the ADDIE model. In addition, we conducted usability test as a formative evaluation.

1. Analysis

1.1 Needs Analysis

The priority of the needs analysis lay in identifying the need for instruction on learning cells and their parts. For this, we assessed both the ideal level and the learners' current level of understanding of cells and their parts. The assessment of the ideal level was achieved through the review of Indiana Academic Standards for science education for fifth and sixth grades, as well as interviews with two elementary school teachers. The standards for fifth grade say that learners must learn that organisms are composed of a single cell or a collection of cells varying in appearance. The standards for sixth grade say that learners must use microscopes to observe cells and recognize cells as the building blocks of all life. Sixth graders are also expected to be able to distinguish the main differences between plant and animal cells. Interviews with the two elementary school teachers (one fifth grade teacher and one sixth grade teacher) confirmed that fifth graders and sixth graders were expected to understand critical characteristics of cells and their parts, such as functions, similarities, and differences of them, with varying focus according to their grade. The two teachers pointed out that the development of supplemental material would be especially useful in assisting learners in understanding the concepts. In addition, any self-paced computer instruction would be useful since there was little time for additional concept practice during class.

For the assessment of the target learners' current level, we administered a questionnaire to a class of twenty eight sixth graders at the elementary school and interviewed five of the students as well as the two teachers. The questionnaire asked three questions on cells and their parts: Can you explain what cells are? Which one of the following (arteries, mitochondria, veins, capillaries) is a cell part? Was it easy to understand cells when you were taught about them? Approximately 65 percent of the respondents failed to give the correct answers to the first two questions. Approximately 61 percent of the students answered no to the third question. Follow-up interviews with five of the students confirmed our findings from the questionnaire that many of our target learners did not have a clear understanding of cells and their parts yet. In addition, the fifth grade teacher mentioned in the interview that students sometimes experienced difficulty in learning so many cell parts and their functions; it can be quite demanding for fifth and sixth graders. She added that though the students succeeded in learning, they tended to end up memorizing the facts without personalizing the information. The needs analysis, therefore, identified a gap between the two levels of understanding of the target concepts; we determined that there was a need for supplemental instruction on the topic in order to bridge the gap.

1.2 Content Analysis

We conducted the content analysis in a series of steps. First, we surveyed science textbooks, teacher guides, science books, and Web sites in order to obtain valid, accurate information. Then we again consulted the fifth grade teacher to obtain useful resources on our topic (such as supplemental materials), and in order to get information on the nature of the content to teach, points to highlight, and teaching strategies appropriate to the grade levels and the topic. Following this, we analyzed the content to teach, primarily based on several established fifth and sixth grade science textbooks and teacher guides. In order to develop accurate and valid content, we crosschecked the science textbooks against each other. In consideration of the nature of the topic and findings of needs and learner analysis, we decided on three sub-topics and detailed content. The content analysis revealed that our instruction actually involved learning twelve different concepts: the cell, the nine cell parts, the plant cell, and the animal cell. These concepts were grouped into three main sub-topics to teach: 1) What are cells?, 2) What are cell parts and their functions?, and 3) Are plant and animal cells the same or different?

Two subject matter experts (SME) reviewed a draft of our instructional material for content accuracy: the fifth grade teacher and a Ph. D. student in science education. They approved the accuracy and completeness of our content with the exception of two points. Based on the two points mentioned by the SMEs, we made corresponding revisions of the content of our instructional materials. After the revisions were made, the 5th grade teacher reviewed our instructional content again and finally confirmed its accuracy and validity. In the process of content analysis, we repeatedly checked to make sure our content was accurate and complete.

1.3 Learner Analysis

We conducted learner analysis through two forms of survey: questionnaire and interview. For convenient data gathering purposes, we used the same survey for both the needs and learner analysis. By looking at the data, we were able to identify the current knowledge of our target learners, determine entry points of our instructional content, and establish appropriate instructional and motivational strategies for our instruction.

The questionnaire consisted of four parts: 1) general information (e.g. gender, level of interest in science), 2) current knowledge of cells and their parts, 3) computer skills, and 4) preferred learning style. Survey data revealed that though all the respondents learned about the topic in their fifth grade, approximately 65 percent of learners failed to provide correct answers to three questions regarding cells and their parts. Thus, we determined that our instruction should begin with explanation of basic concepts of cells as an entry point and extend into more complicated concepts of cell parts. In addition, in response to a question on their preferred way of learning science, students provided varied preferences. The three most common answers were hands-on activities and projects, use of pictures and visual illustrations, and fun ways. We also interviewed five sixth graders to gather more in-depth information on their current science knowledge, interest in science, comfort level of computer skills, study habits, and preferred teaching and learning strategies.

1.4 Context Analysis

Finally, we envisioned the use of our instructional material in various contexts and designed it with an emphasis on flexibility of use. First of all, we decided that our material could be used in conjunction with a classroom lesson as a supplemental tutorial. In this case, the teacher plays a significant role. The teacher will decide to use our material and determine on the exact way of using it. Most typically, the teacher will distribute our entire material to learners and require them to complete it as an assignment. It will be used by all students after a main presentation in a classroom.

Our instructional material could also be used in a completely different context, i.e., independently of the teacher. In this case, parents and learners themselves would decide to use our material. It could be used in a variety of ways, for example, as a refresher lesson for students in need of review or as an enrichment lesson for advanced learners.

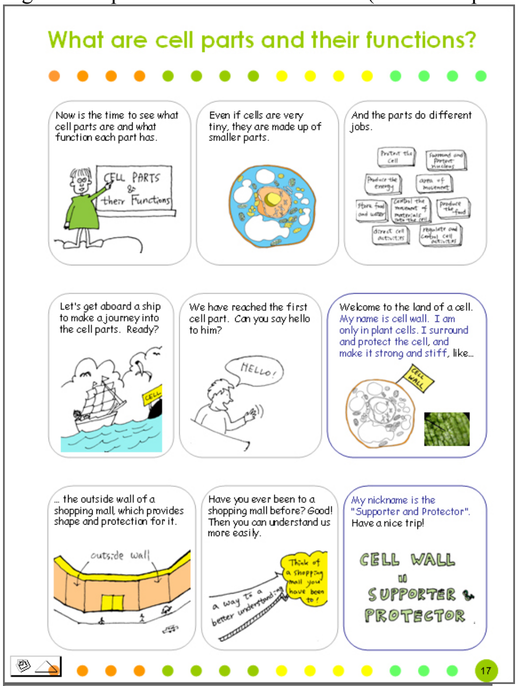
2. Design and Development

2.1 Description of the Instruction

Our instruction is a tutorial composed of learner instructional materials (paper-based and Web-based materials) and a teacher guide. The tutorial may be used flexibly as a supplemental or independent lesson in a classroom or home setting. The learner instructional materials provide the following sections: Introduction, Comic Strips, Things to Remember, Small Practices, Main Practice, and Test. The Main Practice section is Web-based material, whereas the other sections are paper-based. The teacher guide provides instructors with detailed information about how to use the instruction in various contexts.

Among the six sections of learner instructional materials, the Introduction section provides several questions related to learners' prior knowledge and instructional objectives. The Comic Strips section presents three lessons: What are cells?, What are cell parts and their functions?, Are plant and animal cells the same or different?. This section consists of varied forms such as text, narrative, drawings and pictures to help learners to achieve the objectives. Also, this section provides the generality and examples for each concept about cells with the appropriate analogies related to the real world of our learners. After each Comic Strips section, a Things to Remember section summarizes the generality and examples provided in the comic strip with some key sentences and images, in order to facilitate the transfer of new knowledge. After each Things to Remember section, in a Small Practice section, learners answer two or three easy questions about the generalities and examples. The Main Practice section uses the Web to provide classification practice for each concept. The questions in the Main Practice are more difficult than the small practices in the paper-based material. At the end of each question, immediate feedback is automatically given on the learners' performance. Finally, to assess achievement of the instructional objectives, the Test section provides ten yes-no questions. An answer key is provided with the teacher guide. Depending on the use of the tutorial, the teacher may provide the answer key to learners for feedback, or may require the test to be submitted in class. If a parent or student is using the tutorial for independent study, they can likewise decide how to best

Figure 1. Paper-based Learner Materials (Comic Strips and Things to Remember)



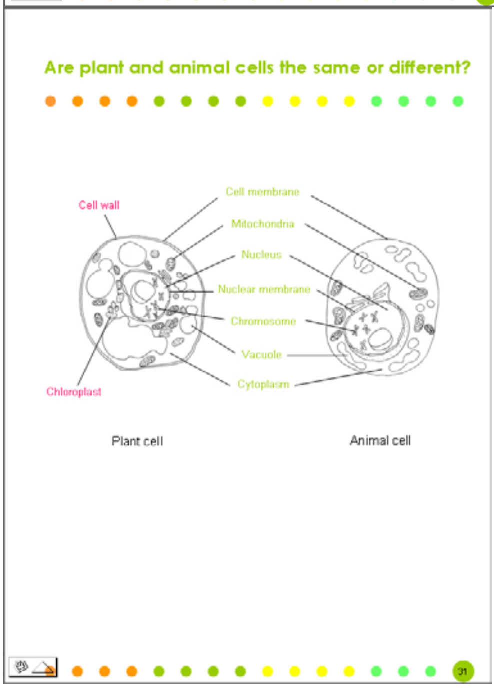
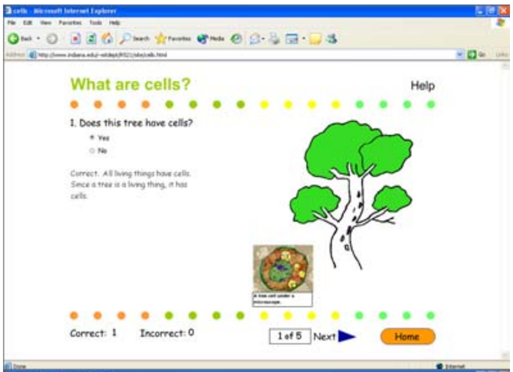
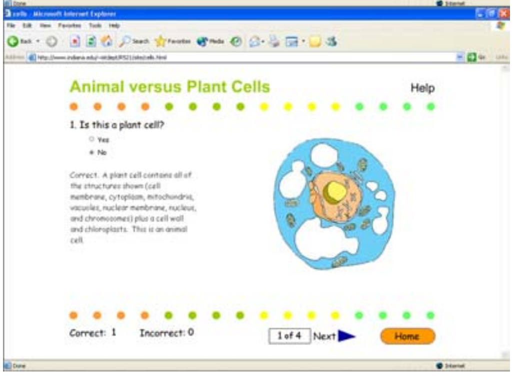


Figure 2. Web-based Learner Material

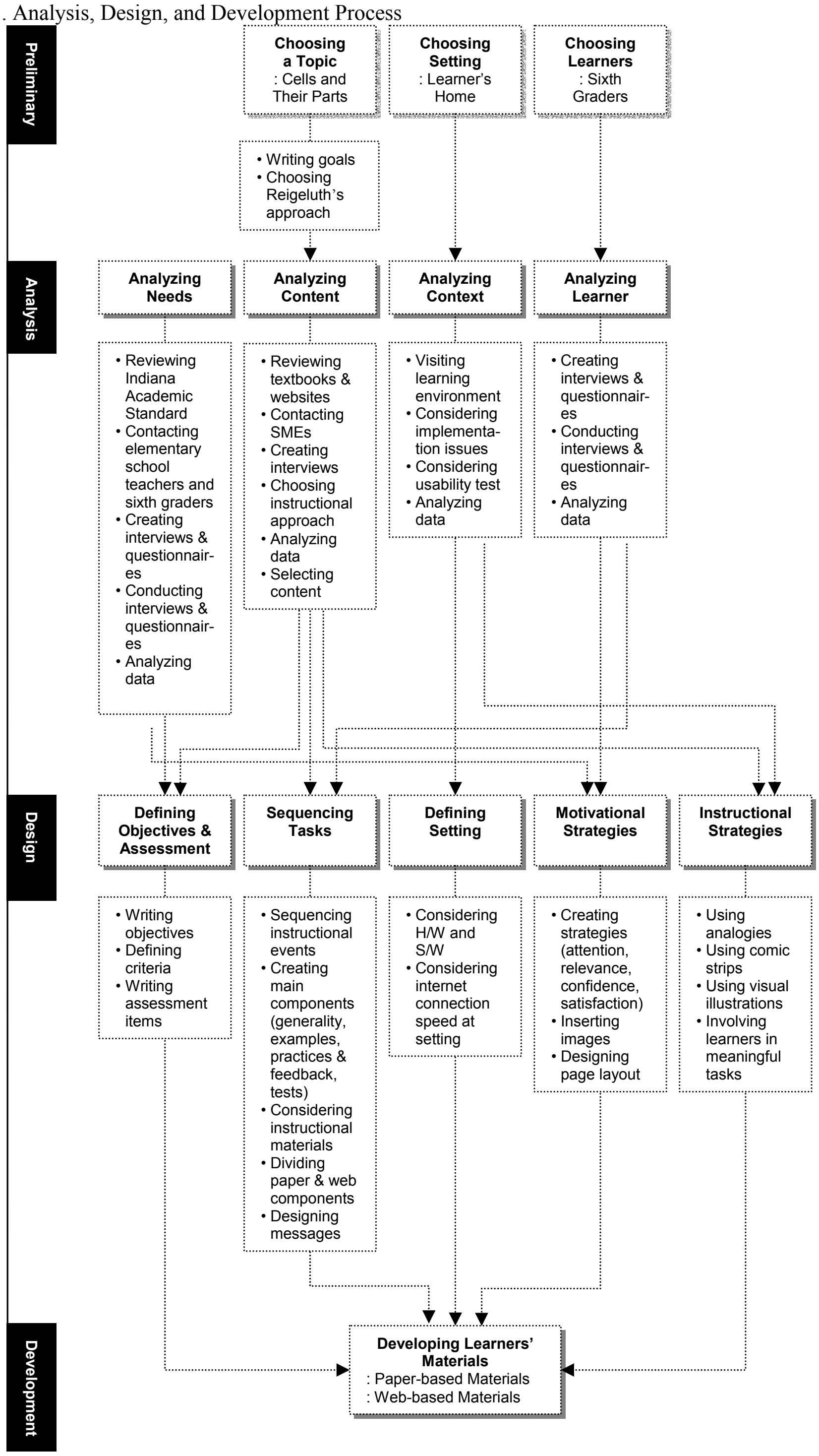




2.2 Process

To design and develop our instruction, we followed the basic steps of the ADDIE model for analysis, design, and development, making some modifications based on this particular project. Figure 1 shows major decisions we made, factors and processes that influenced those decisions, and the sequence of the decision-making. The outputs of one major step (e.g. preliminary, analysis, design, and development) serve as the inputs to the next major step.

Figure 3. Analysis, Design, and Development Process



2.3. Key Design and Development Decisions

For instructional and motivational strategies, we made several key design and development decisions.

The use of paper-based materials for presentation and small practices: We chose to use paper-based materials for these components because the use of paper materials allows learners to keep the materials for future reference and also to write on the materials for the small practices.

The use of Web-based materials for the main practice: We chose to use the Web for the main practice to allow the inclusion of a large number of practice items since practice is important in the learning of concepts. The use of the Web allowed us to add numerous practice items without adding many pages to our paper-based materials. The Web also allowed for the use of instant feedback, as well as easily facilitated repetition of the practice by the student.

The usage of comic strips for learners' motivation: Motivational strategies are especially important for learners to complete instruction, especially when the instruction is being used in a home setting or as an independent lesson. Based on the interviews with target audiences, we decided to present the generality and examples about cells and their parts in the form of comic strips. We expected that the comic strips could gain and keep learner attention so that they would complete our instruction. Also, we anticipated that the familiar stories and narrations of comic strips based on spoken English, not written English, would enable learners to understand the main concepts of the instruction more easily.

The usage of analogies for easy understanding: The name and main functions of each part of a cell still may be difficult for fifth and sixth graders to understand and remember even though we designed them in comic strips to facilitate learner understanding. Therefore, we determined to use appropriate analogies and nick names to explain the name and main functions for each part of a cell as a generality. Also, we used concrete examples related to the real world of fifth and sixth graders to provide familiar examples. For instance, in order to teach the function of a cell membrane, we explained that its function is to hold and protect the cell by using the analogy of an entrance to a shopping mall. We expected the analogies relevance to learners' life environments will motivate learners to study our instruction.

The usage of 'Things to Remember' section as summary: In order to assist learners in summarizing and recalling the generality and examples in comic strips, we inserted the section "Things to Remember" after each comic lesson. This section helps learners to re-organize the content of our instruction in a formal manner, similar to learner textbooks. The key words, images, charts, and diagrams in the section provide a big picture of the lesson, serving as a post-organizer to facilitate the transfer of new knowledge.

The separation of practices into small and main parts: For the practice section, we determined to use the Web-based material to provide more immediate and automatic feedback rather than the paper-based material. However, because of the limitation of short-term memory, learners might forget most content of the lessons if they would get only one change to practice -- at the end of all lessons in the web-based environment. We decided that learners needed some simple practice between lessons and before accessing the Web-based material. We expected that the small practices would engage the learner with the content as he or she progressed through the lessons, and would increase learner confidence by successfully answering some easy questions. This would motivate them for the challenge of the main practice in the Web-based material.

The usage of two types of cell images: line drawings and photographs: To present generality and examples about cells with visualized resources, we decided to use two types of images: line drawings and photographs. As the first presentation in the lesson, we used line drawings to provide the proper detail level that learners need to focus on without including details they don't need. Also, the line drawings allowed us to control color on the points of images we wanted to emphasize. Fifth and sixth graders might have difficulties understanding the main parts of a cell if only real photographs of cells were presented to them by complicating structures of a real cell. Therefore, only after showing the line drawings did we provide photographs as real representation of cells. The real photographs would improve learners' abilities to discriminate real cells and their parts by noting critical and variable characteristics, thus applying their new knowledge to the real world.

The choice of color: orange, yellow, green, and blue: After researching the favorite colors of fifth and sixth graders through the textbooks, we chose four main colors for our instruction: orange, yellow, green, and blue. We expected these colors would motivate learners and satisfy boys and girls' sense of color equally. We properly and consistently used colors to get and focus attention and to emphasize parts in complicated images and large amounts of text in comic strips. Also, we used the same color of cell parts in line drawing to that of the images of the analogies to help learners understand the relationships between two parts. To emphasize important content in comic lessons, we colored the outline of the boxes containing the key content in blue with blue text for the content inside of the boxes.

The consistency of page design, shape, color, and font: To help users to search for information easily and fast, we maintained consistency in the page design, shape, color, and text font. We designed the page layout of our material considering proximity, alignment, and repetition to organize information in an effective way. To enable learners to imagine about cells, the topic of our instruction, we chose a circle and a round square as unifying themes to symbolize cells. We consistently used these themes in all pages of our material. Also we tried to keep the consistent tone of our four main colors when we emphasized the key information of images and text. For example, blue was used for important text and boxes, yellow for names of cell parts, and green for nick names of the cell parts in comic strips. For easy reading, we used only two fonts in all material, Comic Sans MS for comic strips and Arial for the general information. The consistency of these elements in design also was kept in the Web-based material.

The use of Macromedia Flash as development and delivery tool for the main practice: The Flash tool allowed us to create all of the functionality we wanted to include in the practice material. In addition, we also used Flash based on the ubiquity of its plug-in player in most browser software. The platform independence of the tool also allowed us to develop the practice material while allowing users on many different types of computer systems to experience it without browser compatibility issues. This consequently resulted in less user technical problems.

3. Formative Evaluation

3.1. Methods Used to Evaluate the Instruction

Formative evaluation of the instruction occurred during five expert reviews, several informal design team reviews, and a small group usability test. Only the usability test is discussed here.

3.2. Usability Test

The design team held a formal usability test on December 7, 2002. The usability test was conducted at the home of one of the design team members. The room had been arranged with computers for students to complete the computer-based practice.

Since the material had been developed as primarily a self-study tutorial, there was not an instructor present at the usability test. Learners were told that the materials had been designed to be used as self-study and that they should proceed as such.

Five sixth grade learners were invited to participate in the usability test. The learners were selected to represent different ethnicities, fluencies in English, and levels of academic achievement, based upon recommendations by one of the sixth grade teachers that had participated in the analysis phase. On the day of the usability test, however, only three learners were able to participate. Of the three learners present, two learners were Koreans, and the other was Thai-American. The learners had been in the United States from 2-4 years and were fluent in English. The three learners present were above average in their academic achievement level.

In addition to a pre-test, the instruction itself, and a posttest, the usability test included an evaluation questionnaire that asked learners to rate their perceived understanding of the material, the usability of the material, and their satisfaction with the instruction. The evaluation questionnaire also included open-ended questions asking the learners to comment on the strengths and weaknesses of the content, activities, materials, setting, and overall instruction.

3.3. Revision Process

Frick and Boling (2002) recommend a formal process, based on the findings from the usability test and other reviews, to determine the correct interpretation of the data and to decide upon prioritization for changes to be made to the instruction. They recommend prioritizing possible changes based upon importance to the final product and amount of work needed to make the revisions.

The design team followed a similar, though less formal, process in order to determine which revisions should be made. After a review of the all of the collected data and an interpretation agreed upon by the design team members, a final list of possible revisions was compiled and prioritized.

3.4. Revisions Made

The highest priority based on the usability testing was to improve the navigation of the computer-based practice, as two learners had been confused during the usability test. The interface was re-designed to make the "next" and "back" buttons more apparent by including textual as well as graphic information.

A second level priority was a re-consideration of the instruction length as the usability test revealed that students were finishing the instruction much more quickly that the design team had estimated. After several discussions, it was decided that it was not critical that the instruction be completed in a certain amount of time, due to its self-study nature. (Another factor was that since the usability test learners were above average in level of academic achievement, this might have resulted in the shortened completion time.) The important point was to give teachers an estimation of how long a learner might take should the teacher want to assign the instruction as homework or even complete it in class. Thus, the decision was made not to add any additional material to the instruction to make it fill 60 minutes. Rather the teacher and learner materials were slightly edited to include an estimated range of 30 to 60 minutes for completion time.

Summary

Each of the three main components of our theoretical framework supported our instructional design process. The ADDIE model's emphasis on systematic inputs and outputs validated our early analysis and design decisions in the later stages of development. Reigeluth's principles provided a research-based theory for our teaching of concept classification. The use of comic strips aroused initial interest and maintained motivation throughout the lesson. The combination of these factors resulted in effective and appealing instruction for our target learners, as evidenced by the usability test. The project did not, of course, address the implementation and evaluation steps of ADDIE; completing these steps would no doubt provide further information about needed improvements and revisions. Our final process, as described here, however, provides valuable insight that instructional designers and material developers can immediately incorporate into their work.

References

Andree, D. (2001). PS: Happy birthday. *Soldiers,* (June), p. 30-31.

Boling, E. (2002). *R521 course syllabus*. Retrieved October 10, 2003 from Indiana University, Department of Instructional Systems Technology Web site: http://www.indiana.edu/~istdept/R521/syllabus.html.

Carrier, D. (2001). *The aesthetics of comics.* University Park: Pennsylvania State University.

Dick, W., Carey, L., & Carey, J. (2001). Introduction to instructional design. In "The Systematic Design of Instruction" (5th ed.). NY: Longman.

Eisner, W. (1994). *Comics and sequential art.* Tamarac, Florida: Poorhouse Publishing.

Eisner, W. (1996). *Graphic storytelling.* Tamarac, Florida: Poorhouse Publishing.

Fleming, M., & Levie, W.H. (1993). *Instructional message design: Principles from the behavioral and cognitive sciences.* (2nd ed.). Englewood Cliffs, New Jersey: Educational Technology Publications.

Frick, T., & Boling, E. (2002). *Effective web instruction: Handbook for an inquiry-based process.* Unpublished manuscript.

Houghton, H.A & Willows, D.A., (Eds.). (1987). *The psychology of illustration: instructional issues.* New York: Springer-Verlag.

Keller, J.M. (1983). *Motivational design of instruction.* In C.M. Reigeluth (Ed.) Instructional Design Theories and Models: An Overview of their Current Status (pp. 386-429). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Mager, R.F. (1997). Preparing instructional objectives. (3rd ed.). Atlanta, GA: Center for Effective Performance.

McCloud, S. (1994). *Understanding comics.* New York: Harper Collins Publishers.

Misanchuk, E., Boling, E., & Schwier, R. (1999). Visual design for instructional media. Saskatchewan, Canada: University of Saskatoon.

Reigelth, Charles M. (1999). *Basic methods of instruction*. Retrieved October 10, 2003 from http://www.indiana.edu/~idtheory/methods/methods.html

Designing Electronic Collaborative Learning Environments

Paul Kirschner
Jan-Willem Strijbos
Karel Kreijns
Pieter Jelle Beers
Open University of the Netherlands

Abstract

Current research and design of collaborative learning environments - often referred to as Computer Supported Collaborative Learning (CSCL) Environments - tends to focus on surface level characteristics. Educational researchers and designers are busy, for example, determining optimal group size for problem-based education as opposed to project-centered learning. To determine optimal group size, students' collaborative efforts and the results of these efforts are compared for groups of varying size in the different educational settings. This approach resembles comparative research on the use of different media in education that was strongly - and we had hoped definitively - criticized by Clark (1983). He eloquently argued that researchers tend to focus on the media used and surface characteristics of the education they provide. As a consequence, comparative research tends to be inconclusive and the learning materials developed tend to be unreliable at best and mathemathantic (from the Greek: mathema=learning + thanatos=death) at worst. This surface level approach disavows the fundamental differences between the real determinants of learning and behavior in education.

A second problem is that educational institutions tend to apply traditional classroom ideas and pedagogy in non-contiguous collaborative learning environments, assuming that since these environments allow the interaction that we see in the classroom (e.g., chat, real-time meetings, shared applications) traditional pedagogy can be used. The proximate result is often disgruntled or disappointed students and instructors, motivation that is quickly extinguished, poorly used environments, wasted time and money, and showcase environments that are often not much more than computer assisted page turning. The ultimate result is the same as by the first problem, the death of learning.

The solution is as simple as it is elegant, attending not only to technology, but also to the educational and social prerequisites for collaboration. This article provides a framework for designing such collaborative environments based upon the three prerequisites. It then goes into somewhat greater dept with respect to three non-surface level educational factors central to collaboration, namely task ownership, task character and task control. Finally, it presents empirical research on the affordances

Affording collaboration

Instructional design is deterministic or *causal* in that it tends to focus on individual learning outcomes and tries to control instructional variables to create a learning environment that supports the acquisition of a specific skill (a certain person will acquire a specific skill through the implementation of a chosen learning method). With collaboration, the use of groups complicates this. Here, individual and group level variables affect the collaborative learning process such that it is nearly impossible to pre-define conditions of learning or instruction for a group setting such that interaction and skill acquisition are controlled.

Instead of a classic causal view, design of collaborative settings requires a more *probabilistic* view on design. This distinction corresponds with the one made by Van Merriënboer and Kirschner (2001) between the 'World of Knowledge' (outcome) and the 'World of Learning' (process). In the world of knowledge, designers construct methods by which given learning goals in a specific subject matter domain can be attained by the learner. In the world of learning, designers focus on methods supporting learning processes, and not so much attainment of pre-defined goals.

This probabilistic view implies that more attention needs to be paid to learning and interaction processes. Due to the interactions between learners, each person in a group may acquire a given skill through the chosen method, but may equally likely acquire only a part of the skill or the skill and something unforeseen. It might even be the case that the chosen method is abandoned by the group and replaced by another, more idiosyncratic method for that group. The question is not what specific educational techniques and collaborative work forms *cause*, but rather what they actually *afford*, also often referred to as the *affordances* of a learning environment.

What now are these *affordances*? Simply stated, they are the perceived and actual fundamental properties of a thing that determine how the thing could possibly be used (Norman, 1988, p. 9). Affordances are most visible in real life. Some door handles, for example, look like they should be pulled. Their shape leads our brains to believe that is the best way to use them. Other handles look like they should be pushed, a feature often indicated by a bar spanning the width of the door or even a flat plate on the side. Gibson (1977) originally proposed the concept of affordances to refer to the relationship between an object's physical properties and the characteristics of an actor (user) that enables particular interactions between actor and object. "The affordance of anything is a specific combination of the properties of its substance and its surfaces with reference to an animal" (Gibson, 1977, p. 67). These properties/artifacts interact with potential users and *may* provide strong clues as to their operation. In our view, the concept of affordances offers an alternative framework for designing and evaluating IECLEs if appropriated to the educational context.

Education is always a unique combination of technological, social, and educational contexts. Take typical classical education and group learning. The educational contexts are competitive vs. collaborative, the social contexts are individual vs. group, and the technological (physical) contexts are individual workspaces with minimal assortment of materials vs. group workspace with a rich assortment of materials. CSCL represents yet another learning situation. The educational context is collaborative, the social context is the group, and the technological context is a computer-mediated one. The Open University of the Netherlands, for example, uses a computer-mediated communication environment (technological) for competence-based learning grounded in social constructivism (educational) with minimal direct contact, maximal guided individual study, and primarily asynchronous, text based contact between students (social).

When technology mediates the social and educational contexts such that their properties induce and allow specific learning behaviors we speak of 'technology affording learning and education'. This means that we must hold count with technological, social, and educational affordances.

Technological Affordances

Norman (1988) related perceived affordances to the design aspects of an object suggesting how it should be used: "Design is about [real and perceived affordances], but the perceived affordances are what determine usability" (Norman, 1998, p. 123). Norman links affordances to an object's usability, and thus these affordances are designated technological or technology affordances (Gaver, 1991). *Usability* is a well known objective of industrial or product design dealing with physical objects ranging from video-recorders to teapots, and human-computer interaction (HCI) dealing predominantly with *graphical user interfaces* composed of interface objects such as buttons and scrollbars. It is concerned with whether a system allows for the accomplishment of a set of tasks in an efficient and effective way that satisfies the user (e.g., Preece, Rogers, Sharp, Benyon, Holland, & Carey, 1994). It is not a single dimension, but deals with ease of learning, efficiency of use, memorability, error frequency and severity, and subjective satisfaction (Shneiderman, 1998). When creating such environments it is, therefore, important to consider these aspects. Otherwise we risk creating IECLEs that contain all the needed educational and social functionality (in Nielsen's (1994) terms '*utility*'), but that cannot be handled by their users (i.e., the learners) because they are difficult to learn, access, and/or control in the same way video-recorders are.

Social Affordances

Kreijns, Kirschner, and Jochems (2002) define social affordances - analogous to technological affordances - as the "properties of a CSCL environment that act as social-contextual facilitators relevant for the learner's social interaction" (p. 13). Objects that are part of the environment can realize these properties; hence they are designated social affordances devices. When social affordances are perceptible, they invite learners to engage in activities that are in accordance with these affordances, that is, there is social interaction.

In the 'physical' world, affordances abound for casual and inadvertent interactions. In the 'virtual' world, social affordances must be planned and must encompass two relationships. First, there must be a *reciprocal relationship* between group-members and the environment. The environment must fulfill the social intentions of members as soon as these intentions crop up while the social affordances must be meaningful and support or anticipate those social intentions. Second, there must be a *perception-action coupling*. Once a group-member becomes salient (*perception*), the social affordances will not only invite, but will also guide another member to initiate a communication episode (*action*) with the salient member. Salience depends upon factors such as expectations, focus of attention, and/or current context of the fellow member.

Educational Affordances

Kirschner (2002) defines *educational affordances* as those characteristics of an artifact that determine if and how a particular learning behavior could possibly be enacted within a given context. In other words, the chosen educational paradigm – the artifact – is instrumental in determining if and how individual and team learning can take place. Educational affordances can be defined - analogous to social affordances - as the relationships between the properties of an educational intervention and the characteristics of the learners - in the case of environment the learners and the learning groups - that enable particular kinds of learning by them and in the environment for the other members of the group.

Educational affordances in distributed learning groups encompass the same two relationships as social affordances. The environment must fulfill the learning intentions of the member as soon as these intentions crop up while the affordances must be meaningful and must support or anticipate the learning intentions of the group-member. Further, once a learning need becomes salient (perception), the educational affordances will not only invite but will also guide her/him to make use of a learning intervention to satisfy that need (action). The salience of the learning intervention may depend upon factors such as expectations, prior experiences, and/or focus of attention. In the next section we will discuss how these ideas can be incorporated into the design process.

Design Guidelines

According to Norman (1992), the major problem with most new technological devices and programs - and in our opinion also in their use in education - "is that they are badly conceived, developed solely with the goal of using technology. They ignore completely the human side, the needs and the abilities of people who will presumably use the devices" (p. 65). Take, for example, the escalator. This device, a moving stairway, was designed with the intention of speeding up human traffic in a stairway by increasing a person's stair climbing speed with the speed of the escalator. As we all have probably noticed – with the exception of the British – is that it has actually achieved the opposite. Large crowds tend to gather and cue at the top or bottom of the escalator because the users tend to stand still on the escalator itself. This could be the result of human nature (i.e., inherent laziness) that was not entered into the design equation, the non-ergonomic step size that we encounter on almost all escalators (as we all have noticed when we have had to climb or descend an escalator that was standing still) or a combination of the two. Good use – and that means both usability and usefulness - requires a design process grounded in user-centered design research.

We propose a Six-Stage Model for the design of environments on a general level (see Figure 1). In this model, the designer must:

1. *Determine what learners actually do*: Watch students interact, observe collaborating groups interacting to solve problems, observe users interacting with software, et cetera, and do this *before* designing and developing.
2. *Determine what can be done to support those learners*: Determine, based on stage 1, what actually needs to be supported / afforded and then proceed.
3. *Determine the constraints of the learner, learning situation and learning environment and the conventions that already exist*: Look further than the technological constraints and conventions and take into account the educational and social constraints and conventions that play a role in collaborative environments.
4. *Determine how learners perceive and experience the support provided*: There is a world of difference between (good) intentions and user perceptions thereof. Research and design must be carried out as iterative, interacting processes. New 'products' must be tried out with intended users at stages in their development where physical and conceptual changes can still be made.
5. *Determine how the learner actually uses the support provided*: Analogous to stage 1, and following up the more formative evaluations carried out in stage 4 determine if the learner actually does what is hoped or expected.
6. *Determine what has been learnt*: The goal of education is learning and there are three standards to determine the success of any instructional design, namely its effectiveness, its efficiency and the satisfaction of those learning and/or teaching. An increase in one or more of the standards without a concomitant decrease in any of the others means success. This is the proof of the pudding.

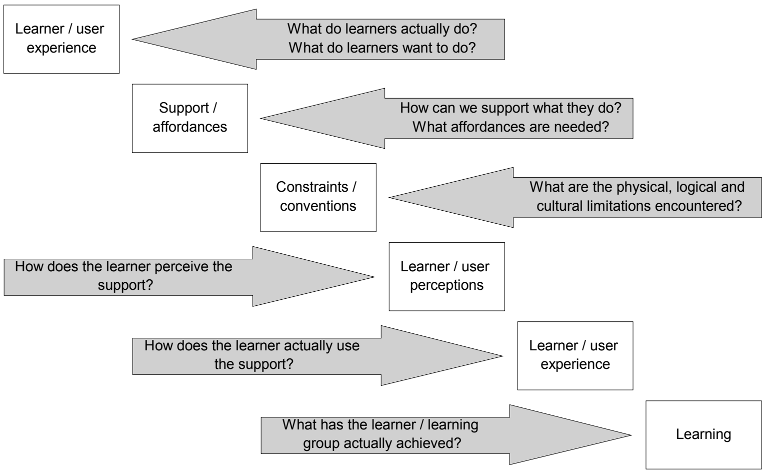


Figure 1. Six-Stage Model of Design

These six stages provide a general approach to instructional design of environments. But, this design needs to also assure that the type of social interaction regarded to be supportive for competency development actually occurs. Thus, complementary to the Six-Stage Model, a more *specific,* process-oriented design methodology is needed which supplies the designer with those questions which must be answered in Stage 3 of the general design level (i.e., to determine the constraints of the learner, learning situation and learning environment). Process-oriented methods may stimulate designers to adopt a probabilistic approach to CSCL design according to the expected interaction, paying attention to critical elements (constraints) affecting the interaction. Strijbos, Martens, and Jochems (in press), propose just such a methodology also consisting of six steps, namely: (1) determining the competencies to be developed, (2) determining the expected (changes in) interaction, (3) selecting the task type, (4) determining whether and how much pre-structuring support is needed, (5) determining the group size which will ensure that the type of social interaction needed for competency development develops, and (6) determining how computer support can be best applied.

Space does not allow a comprehensive operationalization of the two frameworks for the whole of the educational process and all of the actors taking part in it. As most educational design centers around the task, the operationalization of the framework here will focus on Step 3 of the specific design level, i.e. selecting the task-type. We identify three constraints that need to be considered, namely task ownership, task character and task control. We regard *task ownership*, *task character*, and *task control* as defining factors in the educational affording of environments, which will be illustrated through specific prototypical design questions related to these factors. In this discussion we make use of those specific questions pertaining to the third stage of the general model in that they can be used to determine the constraints of the learning environment.

Task Ownership

Ownership in a group is influenced by two pedagogical principles, namely individual accountability and positive independence. *Individual accountability* (Slavin, 1980) was introduced to counter the *free rider* or *hitchhiking* effect: some students would not invest any (or only a minimum of) effort into group performance. By stressing individual accountability, what the group does as a whole becomes less important. It is perfectly valid that in a group environment, each group member is individually accountable for his or her own work. For example, in many problem based learning environments students' sense of individual ownership is increased through grading them for their individual effort, irrespective of the group's performance. *Positive interdependence* (Johnson, 1981) reflects the level to which group members are dependent upon each other for effective group performance (enhanced intra-group interaction). Positive interdependence holds that each individual can be held individually accountable for the work of the group and the group as a whole is responsible for the learning of each of the individual group members. Essential here is social cohesion and a heightened sense of 'belonging' to a group. Positive interdependence is evident when group members in a project-centered learning environment carry out different tasks within a group project, all of which are needed in the final product. Interdependence can be stimulated through the tasks, resources, goals, rewards, roles or the environment itself (Brush, 1998).

Positive interdependence, in turn, provides the context within which *promotive interaction* takes place.

According to Johnson and Johnson (1996), promotive interaction “exists when individuals encourage and facilitate each other's efforts to complete tasks in order to reach the group's goals.” (p. 1028). In other words, individual accountability, positive interdependence, and promotive interaction counter the tendency towards hiding and anonymity.

Task Character

Traditional school tasks are *highly constructed*, oriented towards the individual, and are usually well-structured, short in length, well-defined, and designed to best fit the content and not reality. An archetypical problem is: “Two trains traveling in opposite directions …” On the other side of the coin are *real life* (authentic) problems that are almost always ill-structured (Mitroff, Mason, & Bonoma, 1976) and/or wicked (introduced by Rittel & Weber, 1984; see also Conklin & Weil, 1997). They are often so complex and multifaceted that they can only be solved by multidisciplinary groups working together, where the group members assuage cognitive conflict, elaborate on each other’s contributions and co-construct shared representations and meaning. Examples of such wicked problems are the building a new type of automobile, the legalization of marijuana, and so further.

Complex ill-structured problems require a different educational approach than simple, well-defined ones. To educate people to be able to solve these types of problems we need to choose for a whole task approach, since, after all, real life tasks are not neat segments of some idealized whole. Whole tasks need to be divisible into, non-trivial, authentic part-tasks, which aim at achieving ‘epistemic fluency’. Morrison and Collins (1996) define ‘epistemic fluency’ as “the ability to identify and use different ways of knowing, to understand their different forms of expression and evaluation, and to take the perspective of others who are operating within a different epistemic framework” (p. 109). Ohlsson (1996) enumerates seven epistemic tasks that can be used as guiding principles for the design of collaborative environments (describing, explaining, predicting, arguing, critiquing, explicating, defining), since they indicate the ‘discourse-bound’ activities that learners will have to fulfill during collaborative learning and working.

Task Control

Task control is strongly related to "learner control" which has had a somewhat fluid and eclectic history. In its broadest sense, learner control is the degree to which a learner can direct his/her own learning experience (Shyu & Brown, 1992). More specifically, learner control is the degree to which individuals control the path, pace, and/or contingencies of instruction (Hannafin, 1984). New learning paradigms and new technologies expand this concept of control since they make it possible to provide learners with control over depth of study, range of content, number and type of delivery media, and time spent on learning. With these options, learners can tailor the learning experience to meet their specific needs and interests. For this reason, learner control is not "a unitary construct, but rather a collection of strategies that function in different ways depending upon what is being controlled by whom" (Ross & Morrison, 1989, p. 29). Indeed, learner control may be a continuum of instructional strategies in which the learner is provided with the option for controlling one or more of the parameters of the learning environment (Parsons, 1991). This control can be related to such aspects as context, content, sequencing, pacing, feedback, reinforcement and possibly even learning or presentation style. However, Reeves (1993) points out that a problem with researching learner control centers around what learner control really is, because it could be the pace of learning, as well as sequencing, content, and speed of a program; and much more.

Task control in environments relates to a number of interacting aspects of the total environment that deal with determining the relevant set of actions that students can, should or must perform as well as what an adequate, applicable or best solution or solution path is. It relates to the roles of the teacher/coach versus those of the learners with respect to selecting the relevant activities and learning approach. An underlying assumption here is that learners are amply self-sufficient to be given control over their own learning activities and collaboration methods. Table 4 provides questions relating to design decisions about task control (in terms of ‘pre-structuring’ needed) and the type of competence development that can be afforded.

Three Research Examples of Environments at the Open University of the Netherlands

Functional Roles

CSCL-environments at the Open University of the Netherlands are primarily text-based and asynchronous communication is used. Introducing this form of communication and technological tools introduces students to an unfamiliar realm. An approach to supporting such asynchronous coordination is the use of roles (Brush, 1998). Since roles promote group cohesion and responsibility (Forsyth, 1999; Mudrack & Farrell, 1995), they can be used to foster ‘positive interdependence’ and ‘individual accountability’ (task) ownership). The first research project discussed here focuses on the effect of functional roles that provide context independent process support – developed for higher education – on group interaction and specifically the effect on coordination.

With respect to technological affordances it was decided to limit the technology to the primary process being facilitated: communication. In that respect e-mail was sufficient. Since a sense of belonging to a group is essential to team formation, roles provide a social affordance for the development of group cohesion and a sense of responsibility. Finally, the roles provide an educational affordance given the assumption that they decrease coordination in favor of task-focused communication. In addition, the functional roles used are context-independent and thus transferable to other content domains in which project-centered work is conducted.

Method. Fifty-seven students enrolled in a course on 'policy development' and twenty-three students in a course on 'local government' (49 male and 31 female). Age ranged from 23 to 67 years (Mean = 34.4, SD = 9.03). The study has a quasi-experimental random independent groups design and was conducted in a course on 'policy development'. The manipulation involved the introduction of a prescribed role-instruction in half of the groups (R-groups), aimed at promoting the coordination and organization of activities essential for the group project. The other half of the groups was completely self-reliant with respect to organization and coordination of their activities (NR-groups). Each group consisted of four students and throughout the course they communicated electronically by e-mail. Their task was to collaboratively write a policy report containing an advice regarding 'reorganization of local administration', a topical subject in the Netherlands (and a very 'wicked problem'). For a thorough discussion of the results and methodological considerations of this project see Strijbos, Martens, Jochems and Broers (submitted) and Strijbos, Martens, Jochems and Prins (in preparation).

Results. A non-directional Mann-Whitney test revealed no significant differences between conditions with respect to the group grade ($Z = -1.549$, $df = 4$, $p > .05$). Several scales in the self-report questionnaires comprised a single latent variable that was interpreted as 'perceived group efficiency' (PGE). An F-test for homogeneity of variances to investigate the hypothesis of equality of variances in both conditions for the model without random slope ($F = 2.86$, $df = 4$, $p > .10$) and the model with random slope ($F = 5.86$, $df = 4$, $.05 < p < .10$) revealed a tendency with respect to students' awareness of group efficiency. The outcome of the content analysis corroborates this interpretation.

A non-directional Mann-Whitney test revealed a significant difference with respect to 'task social' statements ($Z = 2.121$, $df = 4$, $p < .05$). Students in the role condition contributed more statements that expressed, either a positive or negative, evaluation or attitude in general, towards the group or towards an individual group member. Furthermore, more 'task content' statements were observed in the role condition ($Z = 1.984$, $df = 4$, $p < .05$). However, the assumption that this would be due a decrease in the amount of coordinative statements was not confirmed. In fact, a directional test revealed that the amount of coordinative statements in the role groups also increased ($Z = 1.776$, $df = 4$, $p < .05$; one-sided). Apparently, functional roles stimulated coordination and as a result 'task content' statements increased as well ($r = .73$, $p < .01$).

A comparison between both conditions of dropout during the course ('quit' category) reveals no differences. However, a comparison of the total number of students that did not finish the course ('quit' and 'no credits') shows a significant difference with respect to the distribution ($\chi^2 = 6,118$, $df = 1$, $p < .05$). Eighteen students in the 'non role' condition - compared to eight students in the 'role' condition - failed to finish the course in time.

Conclusion. The outcomes of this study clearly reveal that focusing solely on group performance outcomes (i.e., grades) as indications of successful collaboration is insubstantial evidence. As well as their design, the study of the environment requires triangulation of research methods and different types of data to construct a representation of the collaborative process.

Group Awareness Widgets

The aim of the second research presented here is to create sociable environments that meet as much of the social (psychological) needs of learners through the explicit embedding of social functionality in them using the framework of social affordances. Within the area of human computer interaction and computer-supported collaborative work, researchers have already become aware that such sociable environments are needed by virtual groups (Donath, 1997; Feenberg, 1989). Bly, Harrison, and Irwin (1993) argue, "The smooth integration of casual and task-specific interactions, combined with the ability to meet informally as well as formally, is a critical aspect of productive group work" (p. 29). In contrast, they observe that "Most tools in computer-supported cooperative work (CSCW) are devoted to the computational support of task-specific activities, but support for cooperative work is not complete without considering all aspects of the work group process. When groups are geographically distributed, it is particularly important not to neglect the need for informal interactions, spontaneous conversations, and even general awareness of people and events at other sites" (p. 29). The kind of social affordances we wish to implement focus on stimulation of informal and casual conversations, stimulation of impromptu encounters, and bridging the 'time gap' imposed by asynchronicity. All three aims imply proximity to be an important dimension of social affordances. The first two aims address proximity of place (i.e., spatial proximity) and the latter addresses proximity of time (i.e., temporal proximity) that can be bridged by using 'traces' which introduce a form of history awareness (Kreijns, Kirschner, & Jochems, 2003).

In CSCL-environments, social affordances devices may be operationalized by *group awareness widgets* (GAWs). These widgets are software tools providing *group awareness* to provide virtual *spatial proximity*, *history awareness*, (i.e. the structured collection of all traces) to provide an overview of *temporal proximitie*s, and a set of *communication media* for *perception-action coupling*.

Based upon the social affordances framework, we have implemented a first prototype of a GAW. This prototype is a first step, not all the implications of the framework are implemented; the set of communication media contains the traditional communication media (chat and e-mail) and pictures of the participants are used.

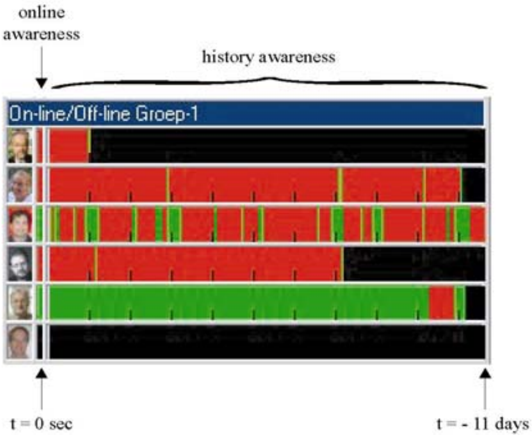


Figure 2. The prototype Group Awareness Widget

The window has a time-axis along which strokes are displayed. The stroke length is an indication of the duration of the engagement. Each member is displayed with her or his strokes. Green (i.e., dark grey if printed) means on-line, red (i.e., also dark grey if printed, but a bit lighter as green) means off-line. Black means that there is no history data available for displaying. As can be seen from figure 2, history awareness is limited to 11 day's; days a separated from each other by small vertical black strokes. Clicking on the portrait of a group member or on a stroke opens a dialog box from which the communication channels are accessible. The dialog box also contains personal and business information about that group member.

As simple as this single segment of the GAW is, it may already invite the group learner to initiate a communication episode that is not based on the learning tasks to be done. It provides the group member with information that may stimulate social interaction with others in the following ways: A group member may perceive that (s)he is not alone in the environment, even in the case that there are no group members currently online; group members may show up at regular times providing opportunities for future contacts in real-time; Patterns of 'busy times' become visible; A real-time conversation can be initiated with group members currently online (i.e., a chance encounter), or a message can be sent to a group member who is currently not online.

In addition to the framework of social affordances, the second research also encompasses the concepts of *sociability* and *social presence* that have a relationship with the concept of social affordances. To this end, instruments are developed that measure the degree of sociability, social space, and social presence (Kreijns, Kirschner, Jochems, & van Buuren, 2003). Testing the hypotheses is the current focus of this research.

A formalism to support common ground

The third study discussed here aims at facilitating grounding in such environments (see Beers, Boshuizen, & Kirschner, 2003 for a more detailed description of this study). The main aim of our research is the design of social/technological affordance that invites and allows users to negotiate meaning while collaborating in context of solving wicked problems. This specific experiment was carried out to explore whether a negotiation widget or tool would positively affect the process of negotiating common ground.

The process of negotiation starts when a team member makes her/his own, as yet unshared knowledge explicit or tangible to others. This can be oral, written, symbolic, et cetera. After one team member has made a contribution, others can try to understand it. In doing this they can consider aspects of the contributor (e.g., educational background, political views, et cetera) as well as their own beliefs and assumptions. A contribution is thus understood against the *presumed perspective of the other*, as well as against one's *own perspective* (Bromme, 2000). According to Clark and Shaefer (1989), the process of grounding continues until "the contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose" (p. 262).

This exploration served as the basis for the design of educational affordances. CSCL environments often make use of external representations formed/restricted by a formalism: a set of objects and rules for making an external representation. In this research a formalism for facilitating negotiation of common ground in problem-solving groups was developed and tested. We expect the formalism to result in more negotiation activities and, ultimately, more common ground. The formalism was embedded in the newsgroup environment by adding specific message types, and specific rules about when one was allowed to post messages of certain types.

Method. In a pilot study we studied people collaborating in a face-to-face setting who were required to making use of a pen-and-paper approximation of the negotiation formalism during their work. Participants were to solve a complex economics problem derived from a dynamic economics simulation game. First, the participants were allowed to practice with the simulation individually, and to solve the case individually, so that they were able to fully apply their own perspective to the task. Next, they solved the problem collaboratively, and after that individually again. All individual problem representations and solutions, as well as the group problem representation and solution, were gathered. The collaboration process was also videotaped.

Procedure. Groups using the formalism were compared with groups not using the formalism. The formalism was built into a widget for a discussion list. A coding scheme for analysis of the collaboration process was developed in which different categories were used for new conversation topics (Contribution), verifying one's own understanding of another's utterance (Verification), clarifying the intended meaning of a previous utterance (Clarification), e.g. as a reaction to a verification, and Elaboration when talk continues about a certain topic, without verifying and clarifying intended meaning of utterances. A high number of clarifications and verifications was seen as indicative for explicit negotiation processes. To get an indication of common ground, overlap between individual problem representations after problem solving was determined. To do this, discussion topics were identified in the coding process to characterise the content of individual representations. Each topic was investigated with respect to whether an individual representation it was present after collaboration. The extent to which specific topics from the discussions were present in more individual representations after collaboration was seen as an indication for common ground.

Results. A preliminary analysis of the data is presented in Table 1. With regard to negotiation, results indicated that the formalism-groups spent more time on negotiation processes than the non-formalism groups (i.e., groups that used their own idiosyncratic representation method), as reflected by the number of utterances that were representative of negotiation. Furthermore, the formalism groups discussed more *different* topics than the idiosyncratic groups. Third, more members of the formalism groups participated in discussion (Participants per segment) than in the idiosyncratic groups. These results suggest a more equal representation of different perspectives in the collaboration process than in the idiosyncratic group. We conclude that the formalism seems to be able to affect negotiation processes by making them more explicit.

	Mean		Standard deviation	
	Formalism	No formalism	Formalism	No formalism
Time (s)	3182	2341	390.58	579.66
Contribution	25.00	19.30	4.36	4.51
Verification	38.00	19.30	5.20	8.50
Clarification	46.00	26.70	7.21	6.66
Elaboration	118.00	102.00	34.51	17.44
Negotiation[a] per contribution	3.37	2.38	.09	.58
Participants per segment	2.67	2.47	.09	.07

[a] Negotiation = negotiation of meaning = sum of verifications and clarifications.

Table 1. *Differences in use of categories*

With regard to common ground, our results indicated that the formalism groups achieved more common ground than the idiosyncratic group (see Table 2). As mentioned, the formalism groups discussed *more* different topics, and captured more of these topics in their group external representation. Furthermore, members of the formalism groups mentioned more different discussion topics in their individual problem representations after the problem solving task, and there appeared to be more agreement between them than between the participants of the idiosyncratic groups (i.e., their individual representations resembled each other more closely than the individual representation made by the members of the idiosyncratic group). We conclude that the formalism indeed seems to have positively influenced the extent to which common ground has been negotiated.

	Mean		Standard deviation	
	Formalism	No formalism	Formalism	No formalism
Total number discussion topics...	25.0	19.3	4.36	4.51
...in the group external representation	19.3	12.0	9.50	5.56
...in the group solution	8.7	2.3	2.52	1.15
...in one individual representation	6.3	6.3	3.06	2.08
...in two individual representations	4.0	2.0	3.00	1.00
...in three individual representations	3.3	2.3	3.08	2.31

Table 2. *Common ground*

These results have to be regarded with some caution. Only six groups were tested, and no statistical tests were used due to our small sample size. The results do provide encouragement to further develop the tool for facilitating the grounding process.

Conclusions

We have outlined a theoretical framework for the design of environments and illustrated this with examples from three research projects. The concept of *affordances* is central to this design and specifically in those cases where the learning environment centers on collaboration. With respect to the design of IECLEs it is not of primary importance what exactly is caused by different elements of the learning environment (learning is no longer causal or deterministic, but has become probabilistic). More important is whether the elements of a learning environment *afford* the type of competency development that was targeted. With respect to collaboration, the question is whether the elements of the environment *afford* the emergence of that type of social interaction that is supportive for the acquisition of the targeted skill.

But these questions cannot be easily answered. We, as designers often think that we know what our designs and products will do and how those, for whom our designs and products are intended, will use them. Unfortunately, this is not always the case. Each of the phases in the design process needs to be studied with respect to the specific choices that can and must be made. Some research is fundamental such as research to determine how IECLE interface design or the way information is presented (e.g., Kester, Kirschner, & van Merriënboer, 2002) affects cognitive load and learning. Other research is more developmental in nature such as research on how group awareness widgets (Kreijns & Kirschner, 2002) affect the perception of sociability, social space and social presence. And still other research is applied such as how specific learners or learner groups perceive a specific IECLE.

Design of CSCL-environments needs to be carried out at two levels, namely a generic and specific level. The impacts of both levels have been illustrated with respect to task ownership, task character and task control. Clearly, the design of a CSCL-environment requires that both levels be taken into account with the specific level being a detailed depiction of the third stage of the general model in that it is used to determine the constraints of the learning environment. Although teachers and designers may prefer a clear set of design rules (i.e., first do A, then do B if you want to achieve C), a deterministic checklist with a limited number of categories is one step too far. We have, however, provided a number of specific design questions to stimulate teachers and designers to think more deeply about their instructional decisions, and not to simply rely on their traditional approach that "has worked so well".

References

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2003, August). *Negotiating shared understanding in collaborative problem solving.* Paper presented at the 10th EARLI biennial meeting, Padova, Italy.

Bly, S., Harrison, S., & Irwin, S. (1993). Media spaces: Bringing people together in a video, audio, and computing environment. *Communications of the ACM*, *36*(1), 28-47.

Bromme, R. (2000). Beyond one's own perspective: The psychology of cognitive interdisciplinarity. In P. Weingart & N. Stehr (Eds.), Practicing interdisciplinarity. Toronto, Canada: University of Toronto Press.

Brush, T. A. (1998). Embedding cooperative learning into the design of integrated learning systems: rationale and guidelines. *Educational Technology, Research & Development, 46*(3), 5-18.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13*, 259-294.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-459.

Conklin, E. J., & Weil, W. (1997) *Wicked problems: Naming the pain in organizations*. Retrieved January 16, 2003 from http://www.touchstone.com/tr/wp/wicked.html

Donath, J. (1997). *Inhabiting the virtual city: The design of social environments for electronic communities*. Unpublished PhD thesis. Cambridge, MA: MIT Press.

Feenberg, A. (1989). The written world: On the theory and practice of computer conferencing. In R. Mason & A. Kaye (Eds.), *Mindwave: Communication, computers and distance education* (pp. 22–39). Oxford: Pergamon Press,

Forsyth, D. R. (1999). Group dynamics (3rd ed.). Belmont, CA: Wadsworth.

Gaver, W. (1991). Technology affordances. *Proceedings of CHI 1991,* (pp. 79-84). New York: ACM.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing* (pp. 67-82). Hillsdale, NJ: Erlbaum.

Hannafin, M. J. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction. *Journal of Instructional Development, 7*(3), 6-10.

Johnson, D. W. (1981). Student-student interaction: the neglected variable in education. *Educational Research, 10*, 5-10.

Johnson, D. W., & Johnson, R. T. (1996). Cooperation and the use of technology. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology*, (pp. 1017-1044). New York: Simon & Schuster Macmillan.

Kirschner, P. (2002). Can we support CSCL? Educational, social and technological affordances for learning. In P. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL*. Inaugural address, Open University of the Netherlands.

Kirschner, P. A., Strijbos, J. W., & Kreijns, K. (2003, in press). Designing integrated electronic collaborative learning environments. In W. Jochems, J. van Merriënboer, & R. Koper (Eds.). *Integrated E-learning: Implications for pedagogy, technology and organization.* London: London: Taylor & Francis Books Ltd.

Kreijns, K., & Kirschner, P. A. (2002). Group awareness widgets for enhancing social interaction in computer-supported collaborative learning environments: Design and implementation. In D. Budny & G. Bjedov (Eds.), *Proceedings of the 32nd ASEE/IEEE Frontiers in Education Conference* (pp. 436-442). Piscataway, NJ: IEEE. Also available at http://fie.engrng.pitt.edu/fie2002/index.htm

Kreijns, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Educational Technology & Society,* 5(1), 8–25.

Kreijns, K. Kirschner, P. A., & Jochems, W. (2003, August). *Supporting social interaction for group dynamics through social affordances in CSCL: Group awareness widgets*. Paper presented at the 10th EARLI biennial meeting, Padova, Italy.

Kreijns, K. Kirschner, P. A., & Jochems, W. (2003, August). *Determining sociability, social space, and social presence in (a)synchronous collaborative groups*. Paper presented at the 10th EARLI biennial meeting, Padova, Italy.

Mitroff, I. I., Mason, R. O., & Bonoma, T. V. (1976). Psychological assumptions, experimentation and real world problems. *Evaluation Quarterly, 2*(4), 639-662.

Morrison, D., & Collins, A. (1996) Epistemic fluency and constructivist learning environments. In B. Wilson (Ed.), *Constructivist learning environments* (pp. 107-119). Englewood Cliffs, NJ: Educational Technology Publications.

Mudrack, P. E., & Farrell, G. M. (1995). An examination of functional role behaviour and its consequences for individuals in group settings. *Small Group Research*, *26*, 542-571.

Nielsen, J. (1994). Usability engineering. San Fransisco, CA: Morgan Kaufmann Publishers (Original work published 1993, Academic Press).

Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.

Norman, D. A. (1992). *Turn signals are the facial expressions of automobiles*. Cambridge, MA: Perseus Publishing.

Norman, D. A. (1998). *The invisible computer*. Cambridge, MA: MIT Press.

Ohlsson, S. (1996). Learning to do and learning to understand. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines* (pp. 37-62). Oxford, UK: Elsevier.

Parsons, J. A. (1991). *A meta-analysis of learner control in computer-based learning environments*. Unpublished doctoral dissertation, Nova University, Fort Lauderdale, Florida, USA.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-computer interaction*. Workingham, UK: Addison-Wesley.

Reeves, T. C. (1993). Pseudoscience in computer-based instruction: The case of learner control research. *Journal of Computer-Based Instruction, 20*(2), 39-46.

Rittel, H. W. J., & Webber, M. M. (1984). Planning problems are wicked problems. In N. Cross (Ed.), *Developments in design methodology* (pp. 135-144). Chichester: John Wiley & Sons. (Originally published as part of Dilemmas in a general theory of planning, *Policy Sciences, 4*, 1973, 155-169).

Ross, S. M., & Morrison, G. R. (1989). In search of a happy medium in instructional technology research: Issues concerning external validity, media replications and learner control. *Educational Technology, Research and Development, 37*(1), 19-33.

Shyu, H. Y., & Brown, S. W. (1992). Learner control versus program control in interactive videodisc instruction: What are the effects in procedural learning? *International Journal of Instructional Media, 19*(2), 85-95.

Slavin, R. E. (1980). Cooperative learning in teams: state of the art. *Educational Psychologist, 15*, 93-111.

Strijbos, J. W., Martens, R. L., & Jochems, W. M. G. (in press). Designing for interaction: six steps to design computer supported group based learning. *Computers & Education*.

Strijbos, J. W., Martens, R. L., Jochems, W. M. G., & Broers, N. J. *The effect of functional roles on group efficiency: Using multilevel modelling and content analysis to investigate computer-supported collaboration in small groups.* Manuscript submitted for publication. Manuscript in preparation for a special issue.

Van Merriënboer, J. J. G., & Kirschner, P. A. (2001). Three worlds of instructional design: state of the art and future directions. *Instructional Science*, *29*, 429-441.

Teacher Education and ICT

Paul Kirschner
Open University of the Netherlands

Abstract

If the Internet is an information superhighway, then teachers just might be the road-kill on the asphalt of the information superhighway (Kirschner & Selinger, 2003). Possibly, for the first time in history, students are more adept at using the tools necessary for acquiring and transmitting knowledge than are their teachers. Children everywhere are creating their own virtual communities through the use of new technologies. They make use of chat facilities (e.g., MSN®, ICQ®) to stay synchronously in touch with both old and new friends and email and short message services (SMS) asynchronously. They take part in discussion groups, navigate through virtual worlds, and assimilate new hardware and software as if it were a second nature. In many ways they are light-years ahead of their parents and teachers regarding the possibilities and use of ICT. As a result students are getting bored and frustrated and teachers are getting frustrated and distraught. This paper presents examples of the pedagogy of good practice and benchmarks for calibration and/or modeling of ICT-teacher training.

State of affairs of teacher education with respect to ICT

Information and Communication Technology (ICT) has the potential to meet the learning needs of individual students; promote equality of opportunity; offer high quality learning materials; and increase self-efficacy and independence of learning amongst students of all ages. For the teaching profession, ICT is not only an essential tool for teachers in their daily work, but it also offers them opportunities for their own professional development. It can be used to encourage new ways of working as part of professional learning teams and it offers schools themselves the possibility of a faster route to establishing a meaningful role in the wider community, embracing learners of all ages, linking and networking to other educational establishments and bringing professionals together across a range of areas.

But all that glitters is apparently not gold. An ICT adviser for a large Local Education Authority in the UK reported that he has observed an increase in the use of ICT but it was just to be 'more of the same - we are seeing nothing transformational'. The same adviser added that he did not believe that learners have more autonomy; technology was not being used to give students a new way into learning, nor was there a change in pedagogical practice. Cuban (1993), critical of the developments in the use of ICT in schools, believes teachers appropriate new technologies and incorporate them into their traditionally held views of teaching and learning. He argues that the overhead projector and video made very little impact on teaching styles, so why should computers be any different? But computers *are* substantially different from any previous technologies because multimedia and hypertext give students access to new ways of thinking through dynamic images, simulations and models, and the Internet provides access to a huge array of previously untapped information. Teachers must find ways of harnessing the power of the new technology. Their jobs will change but their role should become no less important in the same way that public libraries and books did not make teachers redundant.

There is thus a growing tension between demands for radical change in educational priorities and processes and the expectations for teachers, especially with respect to the use of ICT in the teaching and learning process. Whether these priorities are determined locally or nationally, the teaching profession is having trouble adapting to these priorities at a pace that is fast enough to support the radical changes demanded. Teachers are increasingly on the front line in implementing policies designed to reap educational benefits from investments in ICT. Innovations in the effectiveness of teacher recruitment, training and professional development have become key issues and will remain so.

De Haan, Huysmans, and Steyaert (2002) reported that 97% of all Dutch high school students questioned stated having a computer at home, 84% had an Internet connection at home, 80% had their own email address, and 22% even had their own website. They use their computers more at home than at school (16% of the students report not having access to a computer at school compared to only 3% reporting not having access at home) and if they report using computers at all at school, it is not usually in their lessons. For the required course on Informatics, only 41% of the students reported using the computer at least <u>once</u> a month! In another study examining students' views of ICT in school and the home (Mumtaz, 2001) students reported their experiences at school being inferior to their experiences at home. They described school use as boring and cited examples such as 'typing up from the board' as a use of ICT, adding weight to Cuban's perspective.

Given that schools are increasingly expected to integrate technology and provide increased access to students and teachers, these results reveal that a majority of teacher training students are graduating in an information age without proper guidance on how to use technology in the classroom. There is, thus a GAP!

Core technology

ICT is a core technology in the teacher training/learning setting (Collis & Moonen, 2001). They define a core technology as being the major way of organizing the learning experience; the component around which all other components are planned. In contrast, complementary technologies are optional, serving a valuable function but able to be compensated for via the core technology if so needed, or dropped altogether if not functioning or feasible. They also speak of the goals as learning how to use ICT and learning with ICT. In learning *how to use* ICT, the focus of teacher training is how to use such products in the classroom and off-campus. Teachers face new roles with respect to using ICT and must learn how to fulfill those roles. In learning *via* ICT the presentation and distribution of instruction is primarily through web environments or systems offering an integrated range of tools to support learning and communication. The synthesis of these two axes can be seen in Figure 1.

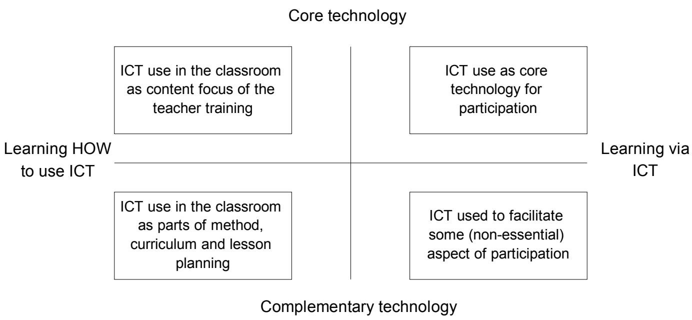


Figure 1. ICT in teacher training

The research reported on in this paper concentrates on ICT as core technology, and thus on the upper two quadrants. The upper-left ICT use as the content of teacher training refers to helping teachers gain competence with ICT, for example, with applications, specific educational software packages and the Internet. The upper-right quadrant relates to ICT (predominately web environments) as the tool used to support flexible learning for teachers, and particularly for specialist or postgraduate school- or home-based study for teachers, just-in-time professional development including networking with other teachers, mentoring new teachers, and inter-regional or international collaboration. The online learning networks for teachers provided in many parts of the world are examples of teacher learning via ICT as a core technology.

Pedagogic benchmarks for ICT teacher education

Many countries and agencies are developing and guiding professional development to help prepare teachers to use ICT in education. UNESCO, for example, has been working to produce guidelines for less developed countries at the request of their governments (UNESCO, 2002). Many developed countries are also promoting initiatives for ICT teacher education. Niki Davis, in her final editorial for the Journal of Information Technology and Teacher Education (2000) noted that ICT in teacher education is at a first zenith and that 'the heat is on' for those who prepare educators for effective use of information and communication technologies.

Davis and Tearle (1998) reviewed frameworks for ICT in teacher education in 1998 to inform a European Commission research objective for a 'core curriculum' for ICT in teacher education (then known as Telematics for Teacher Training). They noted that many countries around the world were taking action to ensure that their educational systems are updated to permit equality of access and to ensure that the key ICT skills are developed in schools and other educational institutions and that it was becoming abundantly clear that the training of teachers in ICT skills and appropriate pedagogical approaches is essential. Davis and Tearle led the creation of a holistic framework to guide good practice (the T3 core curriculum - http://telematics.ex.ac.uk/T3). It was expanded and updates by the UNESCO task force on ICT teacher education during 2001/2 (Davis, 2003).

The Methodology of this study

The research here can be characterized as a quick-scan of initiatives in the field of teacher training across the globe. Five *experts* in the field of ICT and teacher education from around the world did the research as an asynchronous distributed research group which made use of a *web-based project environment* for determining the reference framework, sharing relevant documents, cases, and web sites, for discussion of practices found and the collection of data. After contracting the experts the group as a whole – making use of the document archive and threaded discussion forum, the group developed a *reference framework* based upon Collis and Moonen's (2001) categories for ICT in teacher training. Based upon this reference / study framework, three actions ensued.

First, the experts used their knowledge of the field and their own professional networks to locate examples of good practice. ‘Good practice’ is a series of documented strategies and tactics employed by highly admired institutions trying to find an effective, efficient and satisfying solution for identified needs, through processing and/or prototyping, to a finished product (paraphrase of Arthur Anderson, 2000). These institutions are not ‘best-in-class’ in every area - such institutions do not exist - but due to the nature of competition and a drive for excellence, the profiled practices have been implemented and honed to help place their practitioners as examples of ‘how it can be done’. Second, the team developed two instruments for documenting the practices: a checklist and an evaluation form. Finally, the experts filled in the forms and supplied additional documentation so that the team leader in the Netherlands could begin a lines-of-reasoning synthesis (meta)-analysis.

Results

Twenty-six examples of good practice in the preparation of (student) teachers for working in an ICT-rich environment were collected from five world regions (see appendix 1). The practices were analyzed with respect to the emphasis placed on different aspects of ICT in teacher training, the depth and breadth of the practice(s), and the pedagogy employed. For an in depth discussion of these aspects, the reader is referred to Kirschner and Davis (2003) in a special issue on ICT and Teacher Training of Technology, Pedagogy and Education’s (2003, vol. 12, nr. 1). What follows here are the benchmarks and pedagogic underpinnings of the research (see Table 1).

The pedagogical underpinnings (and there sub-categories) are based upon the pedagogical dimensions of computer-based education as defined by Reeves (1994), which was used in the evaluation form filled in by the reporters. Reeves’ dimensions are based on different aspects of learning theory and can be seen as criteria for evaluating different forms of computer based education (see table 2).

Become a competent personal user of ICT

All cases collected included ‘Becoming a competent user of ICT’. Initial ICT-skills are being taught or their acquisition supported in all cases, but the breadth and depth of application is dependent on the individual scope of the program or course. A gap found with respect to *using ICT as a work tool* - not reported as being a part of 7 of the 26 cases - and as a *communication tool* between the school, parents, local community and society in general – not a part of 6 of the 26 cases - is a little unsettling. Since computer supported collaborative or co-operative learning is one of the carriers of the new constructivist philosophy of learning and education, teachers need to be able to use such tools. If teacher education wants to prepare their students for working in such environments that will often design and develop learning materials in teams, then it would be wise for them not only to ‘talk the talk’ of educational innovation, but also to ‘walk the walk’. Selinger states: "Schools can evolve into community centres to further promote student connections to the community around them and work with groups on real-world projects. The online delivery of education will provide a means to centralize course development and links to academic tutors on a global scale. Email and computer discussion forums are ideal for disseminating good practice between teachers. Although teachers are enthusiastic about these forums, they report that they are unable to make full use of them as the time available to do so is limited." (2001, p.15) Preparation and experience with these tools during ICT teacher education could prepare teachers and reduce the time they would need to spend on gaining skills in their effective use while practicing teachers.

Competently use ICT as a tool for teaching

All but one of the cases identified the objective that students to ‘learn to competently use ICT as a tool for teaching’ (at the general level). Inspection of the specific results shows that within this aspect certain specific goals are (almost) always prevalent while others are not preferred. These are adapting technologies to better teaching, planning and learning material preparation and production, which are seen as important elements in good practices. The following sub-categories were not included for in more than six of the twenty-six cases are given in order of increasing exception:

- Team teaching in situ or at a distance - not included in 7 programs
- Manage learning - not included in 8 programs
- Implications of different machines and platforms - not included in 9 programs
- Managing teaching - not included in 11 programs
- Teaching at a distance - not included in 12 programs
- Instructional design of course materials - not included in 14 programs
- IP telephony, video conferencing - not included in 17 programs

Competently use ICT as a mindtool

Mindtools are computer applications that, when used by learners to represent what they know, necessarily engage them in critical thinking about the content they are studying (Jonassen, 1996). Mindtools scaffold different forms of reasoning about content and require students to think about what they know in different, meaningful ways. Using databases to organize a student's understanding of content organization necessarily engages her/him in analytical reasoning, where creating an expert system rule base requires thinking about the causal relationships between ideas. At this point a distinction need be made between learning *with* ICT (as a mindtool) and learning *from* ICT (as a mindtool). In learning *with* ICT we speak of short-term goals where the ICT is the enabler. Using project-planning software in project-centered learning will - in the short-term - help students plan a project properly and hand in projects on time. Learning *from* ICT involves long-term goals where ICT causes a change in the way one thinks and works. Using the same example, in the long run project-planning software can teach students to organize their thoughts, take critical paths and products into account, and plan the work efficiently (long) after completing the project. This is learning from ICT (and using ICT as a mindtool).

> Learning with mindtools depends "on the mindful engagement of learners in the tasks afforded by these tools and that there is the possibility of qualitatively upgrading the performance of the joint system of learner plus technology." In other words, when students work with computer technologies, instead of being controlled by them, they enhance the capabilities of the computer, and the computer enhances their thinking and learning. The result of an intellectual partnership with the computer is that the whole of learning becomes greater than the sum of its parts" (Jonassen, Carr, & Yueh, 1998, pg. 14).

All programs attempt to train (student) teachers to become competent users of ICT as a mindtool. Many computer applications have been developed explicitly to engage learners in critical thinking while others can be repurposed to do this. Examples of mindtools include semantic organization tools, dynamic modeling tools, information interpretation tools, knowledge construction tools, and conversation and collaboration tools. This last type of mind tool, the conversation tool, is increasingly used. Probably the most important reason for this increase is the adoption in education of theories of learning where both the social and the constructivist nature of learning are emphasized. Software that promotes reflective discourse can also be used as a way of bringing the theoretical aspect of teacher training taught in university or colleges together with the practice experience and articulated in schools in which beginning teachers undertake teaching practice (Pearson & Selinger, 1999). Several aspects of collaboration (with other teachers, experts, designers) and co-operation (between teachers, teacher educators and student teachers) were highlighted. The focus on collaboration with other teachers, experts and designers is part of twenty programs and cooperation between teacher, teacher educators and student teachers forms a component of twenty-one programs.

Master a range of educational/pedagogical paradigms that use ICT

All cases aimed to help students gain a command of a range of educational / pedagogical paradigms which make use of ICT. In their programs of study, learners consider contemporary insights concerning learning and teaching and the role ICT plays within these processes. Selinger (2001) noted that it is often the case that the increase in the use of ICT moves beyond 'more of the same' to education. The CEO forum to inform the development of ICT in schools and teacher education describes progression of schools, colleges and universities as they transform education (http://www.ceoforum.org). However, when learners are not given more autonomy; technology is not used to give students new ways of learning, and there is very little change in pedagogical practice. According to Cuban (1993), teachers tend to appropriate new technologies and incorporate them into their traditionally held views of teaching and learning. Cuban argues that the overhead projector and video made very little impact on teaching styles, so why should computers be any different? But computers are substantially different from any previous technologies because multimedia and hypertext give students access to new ways of thinking through dynamic images, simulations and models, and the Internet provides access to a huge array of previously untapped information. Teachers must adapt and transform their practice in order to find ways of harnessing the power of the new technology. Their jobs will change but their role should become no less important in the same way that public libraries and books did not make teachers redundant.

Chris Yapp (2001) put it the following way: "It is a great fallacy that "New teacher = Old teacher + ICT". Instead, training teachers in the use of ICT in their own curriculum areas should be the start of a process of educational transformation process, not its completion."

Master a range of assessment paradigms that use ICT

New paradigms about learning and teaching must include new views and visions of assessment. Technological developments create new possibilities to what often is called 'alternative assessment' (Reeves, 2000). Teachers and aspiring teachers should be aware of the range of assessment paradigms – including the complex 'alternative' forms of assessment needed for assessing the attainment of higher order educational goals that involve deep understanding and active use of knowledge in complex, realistic contexts that may make use of ICT. Most of the cases (20 out of 26) take note of this standard. Both the limitations and affordances of ICT tools in assessment are present in 19 cases and new ways of assessing learning is seen as important. Several cases incorporate new modes of assessment for students in their programs of study:

Understand the policy dimension of using ICT for teaching/learning

Research and practice provide evidence that ICT implementation at the classroom (micro) level requires policy at the school (meso) and district/national (macro) level. In the UK in 2001, this aspect was a mandatory part of the Teacher Training Agency Standards required of newly qualified teachers. For example secondary teachers have to have, for their specialist subject(s), where applicable, a detailed knowledge and understanding of the National Curriculum programs of study, level descriptions or end of key stage descriptions where applicable. Most states in the USA have also mandated standards for ICT in teacher education, many simply adopting International Society for Technology in Education standards (2001). Although eighteen cases included this ‘policy-standard’, it is not always clear how this element is implemented in these cases.

Nine of the programs do not appear to pay attention to translating government policy to the classroom and an alarming number almost half report that they pay no attention to the critical step in policy implementation, namely *assessing the situation in policy terms*. However, care should be taken in interpreting the data because policy and its implementation in education as sensitive to cultural differences and may change with speed.

Other (often social) aspects of ICT use in education

Almost all of the cases report education in this broad area. These include ICT and citizenship, innovative democratic learning supported by ICT, and using ICT to help integrate education and learning with sport, welfare, community education, and in training and development facilities in industry (hospitality, banking, government departments).

Pedagogical aspects

With respect to the pedagogical aspects, the results can be seen in table 2. The box approximating the average score has been checked. Where the average score lies approximately mid-way two values (e.g., 3.5) the two boxes have been checked.

								M	*SD*
Epistemology	Objectivist	☐	☐	☐	☒	☐	Constructivist	4.13	1.01
Pedagogical philosophy	Instructivist	☐	☐	☐	☒	☐	Constructivist	4.17	.98
Underlying psychology	Behavioural	☐	☐	☐	☒	☐	Cognitive	4.13	.63
Goal orientation	Sharply-focused	☐	☒	☒	☐	☐	Unfocused	2.43	1.16
Experiential validity	Abstract	☐	☐	☐	☒	☐	Concrete	4.26	.69
Contextual validity*	Contextualised	☐	☒	☐	☐	☐	Decontextualised	1.96	.82
Flexibility	Teacher-proof	☐	☐	☐	☒	☐	Easily modifiable	3.96	.77
Individual differences	Non-existent	☐	☐	☐	☒	☒	Multi-faceted	4.48	.51
Learner control	Non-existent	☐	☐	☒	☒	☐	Unrestricted	3.48	.85
User activity	Mathemagenic	☐	☐	☐	☒	☐	Generative	4.19	.60
Co-operative learning	Unsupported	☐	☐	☐	☒	☒	Integral	4.35	.65
Cultural sensitivity	Non-existent	☐	☐	☐	☒	☐	Integral	3.87	.81

* Contextualised = Use 'immediately'; Decontextualised = Learn now, use later

Table 2. Pedagogical aspects of the good practices of ICT teacher education (Reeves, 1994).

Examination of Table 2 shows that several of the 26 cases of good practice analyzed adopted the underlying epistemology, philosophy and psychology. However, it was noted earlier that a range of pedagogies were used and therefore most cases, if not all, incorporated some behaviorist approaches. The goal orientation and contextual validity which are biased towards sharper focus and contextualization also fit with authentic learning about ICT in teacher education that will ease transition to good practice in the workplace. Indeed, many cases incorporate service learning and work to improved education and support of teachers and teacher educators.

The programs chosen conform largely to the ideas behind modern constructivist learning. Constructivism has come to be an umbrella term for a wide diversity of views. Most theories, though, view learning as an active process of constructing rather than acquiring knowledge, and instruction as a process of supporting that construction rather than simply communicating knowledge (Duffy & Cunningham, 1996). Engaging students with challenging real-life contexts where the environment is rich in information and where there are no right answers best situate learning. The recommended tasks are authentic and learning is best approached through ‘cognitive apprenticeship’. Learners negotiate through interactions with others and in this way they can learn to develop the necessary multiple perspectives on reality that are especially important (when learning takes place in ill-structured domains). Reflexivity is essential and must be nurtured. Kirschner (2000) discusses learning in this way using integrated electronic learning environments.

The Benchmarks

Benchmark 1 -Personal ICT-competencies

Programs for teacher training should facilitate aspiring and practicing teachers to be competent personal users of ICT. Minimally, teachers should have or develop basic competencies involving the use of:

- office applications such as word processing, spreadsheets, databases, drawing packages, and a simple web page editor
- resource tools such as CD-ROMs, Internet, Web-portals, different types of search machines
- communication tools such as email, listserv and synchronous chat.

Beyond this, a program for teacher training should develop the learners ability to use ICT effectively for:

- communication between and within student groups
- communication between and with other teachers
- continuing their own education once they have completed their studies, including self-assessment of own learning and learning needs

Benchmark 2 - ICT as a mind tool

Programs for teacher training should train aspiring teachers to be able to make use of ICT as mindtools. As previously discussed, mindtools are computer programs and applications (for example those in Benchmark 1) that facilitate meaningful professional thinking and working. Mindtools help users represent what they know as they transform information into knowledge; they are used to engage in, and facilitate, critical thinking and higher order learning. Minimally, teachers should develop basic competencies to use mindtools for:

- co-operation (between teachers, teacher educators, and student teachers), and
- collaboration on pedagogical projects (with other teachers, experts, designers, etc.)

Benchmark 3 - Educational/pedagogical use of ICT

Programs for teacher training should train aspiring teachers to be able to make use of ICT within many different educational/pedagogical settings. In other words, not in adApting their education to ICT, but rather of adOpting ICT in their education. Minimally, teachers should develop basic competencies to use ICT effectively for:

- collaboration / co-operation in both asynchronous (email, discussion lists, web based forums, listservs) and synchronous (video, audio, chat, whiteboard, file sharing) environments
- resource based learning (informing, asking questions, evaluating, comparing)

Related to this is need for teacher training organizations to deal with the pedagogical uses of ICT at a meso-level, for example for comparing and selecting resources such as:

- learning environments,
- project environments,
- collaborative environments,
- learning management systems, and
- software.

Finally, it is of paramount importance that programs for teacher training acquaint and prepare aspiring teachers and teacher educators themselves with the effects of ICT on:

- their own role as teacher and
- their students to increase autonomy, authentic activity, learning styles, situated learning, motivation, enfranchising those who are out of the mainstream.

Benchmark 4 - ICT as a tool for teaching

The major pitfall that must be avoided when training teachers to use ICT is using the tool for the tool's sake. None of the cases of best practice included this and many modeled innovative practices where ICT enhanced and extended the curriculum and served to develop education. ICT must be used to meet educational objectives in a way that is integrated into the school program. This means that aspiring teachers not only come to know the theory behind why and how to use ICT, but will also develop competencies in:

- adapting technologies TO good/better teaching such that the leaching/learning can change for the better,
- planning for relevant individual, group and whole class activities,
- preparing and producing of learning materials with the help of ICT,
- dealing with the possibilities / consequences of using ICT,
- teaching and learning specialist subject(s) with ICT, and
- team teaching in situ or at a distance.

Benchmark 5 - Social aspects of ICT use in education

ICT is having, and will continue to have, a profound effect on society. Traditional normative concepts such as privacy, anonymity and intimacy are changing. Norms and values have traditionally been passed from adults to children, but now the children are also engaged at the cutting edge of societal change. With instant messengers they multi-task conversations in ways that adults are hard pressed to understand. For example, breaking news a terrorist attack can now often known more quickly on the other side of the world than on the other side of the city. It is important that teachers and teacher educators:

- engage as member of a (wired) school community,
- provide a role model of good ICT practice,
- learn to share and build knowledge,
- understand the implications of the Information Age on schools and schooling, and
- realize an discuss the impact of ICT on society.

Future benchmarks: ICT in assessment and policy

Learning to use ICT for assessment and understanding the policy dimension of ICT-use are not widely perceived as good practice at this time. In the opinion of the author this is shortsighted, especially assessment. Assessment via ICT, and in particular new forms of assessment involving the learner as a collaborator in assessment (e-portfolios, learning diaries), peer assessment, and authentic assessment are of growing importance.

The need for a benchmark on policy is less clear. It is possible that this item was misinterpreted. It would be strange for learners to remain ignorant of local standards regarding ICT in their educational system, especially where ICT is mandatory or integrated into mandatory standards for the subjects the teacher is preparing to teach. For example, most US states and the US Council for the Accreditation of Teacher Education have adopted the ISTE standards (2001) as mandatory. An alternative interpretation of this question is related to the development of ICT policy for education, and this is not a recommended benchmark for all aspiring teachers education in ICT.

Conclusion

Six benchmarks of good practice for both pre- and in-service programs for teacher education were that teachers become:

1. Competent personal users of ICT,
2. Competent to make use of ICT as a mind tool,
3. Master a range of educational paradigms that make use of ICT, and
4. Competent to make use of ICT as a tool for teaching
5. Master a range of assessment paradigms which make use of ICT
6. Understand the policy dimension of the use of ICT for teaching and learning

The first four benchmarks are characteristics found in almost all of the programs. The final two were also seen as important, but were not always present in the cases evaluated. The experts also agreed that, in cases of best practice, the benchmarks would be seen in an environment that not only talks of modern constructivist thinking and pedagogy, but one that also adopts and models those practices. The days of teaching ABOUT the use of ICT are over and directed teaching of ICT skills is not recommended.

> If we want teachers, students, administrators, politicians, instructional designers, et cetera to consider the computer, other information appliances, and the Internet as ‘normal’, then it’s time to stop treating them as something special. We do not speak of, do research on, or specifically teach our future teachers about Book Assisted Instruction (BAI) although our present educational system is almost totally based upon BAI. Teachers can plan their lessons, choose what to ‘teach’, what to expand upon, and what to leave to individual self-study thanks to our system of BAI. Departments can coordinate curricula, students know what to study, and parents know what to coach all thanks to BAI. Testing, certification institutes, and accreditation agencies know what to test, certify, and accredit thanks to schoolbooks … But no one talks about BAI or for that matter about BBAI (blackboard assisted instruction), TAI (teacher assisted instruction), and so forth. They are seen as *normal* parts of education – both teaching and learning – and are treated as such. Why then do we continue to speak of CAI, why do we still have special courses on computers and pedagogy, why…? As long as we continue to think of computers and computer networks as something special in education, they will never become ‘NORMAL’.

References

Arthur Anderson - Contract Services (2000). Control Self-Assessment Technique. Best Practices, LLC. Retrieved on May 2, 2002 from http://www.best-in-class.com/

CEO Forum (2001). CEO Forum on Education & Technology [Online] Retrieved July 3, 2002 from http://www.ceoforum.org

Collis, B., & Moonen, J. (2001). *Flexible learning in a digital world: Experiences and expectations*. London: Kogan Page.

Cuban L .(1993). Computers meet classroom, classroom wins, *Teachers College Record, 95*(2), 185-210.

Davis, N. (2000) Editorial. *Journal of Information Technology for Teacher Education, 9*(2), 277-286.

Davis, N. (2003). Technology and teacher education in the USA: What makes for sustainable good practice? *Technology, Pedagogy and Education, 12*(1)*,* 59-84.

Davis N., Desforges, C., Jessel, J., & Somekh, B. (1997) Can quality in learning be enhanced through the use of IT? In B. Somekh and N. E. Davis (Eds.), *Using IT effectively in teaching and learning: studies in pre-service and in-service teacher education* (pp. 14-27). London: Routledge.

Davis, N. & Tearle, P. (1998). A core curriculum for telematics in teacher training. *Teleteaching 98 Conference*, Vienna, Austria. Retrieved May 2, 2002 from http://www.ex.ac.uk/telematics/T3/corecurr/tteach98.htm

De Haan, J., Huysmans, F., & Steyaert, J. (2002). *Van huis uit digital* [Digital at home]. The Hague: SCP.

Duffy, T. M. & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. Jonassen (Ed.). *Handbook of research for educational communications and technology*. New York: MacMillan.

INTIME (Integrating New Technologies Into the Methods of Education). Retrieved on May 2, 2002 from http://www.intime.uni.edu/model/model.html

International Society for Technology in Education. (2001). Technology standards for teachers. [online] Retrieved on May 2, 2002 from http://www.iste.org

Jonassen, D. H. (1996). *Computers in the classroom: Mindtools for critical thinking*. Columbus, OH: Merrill/Prentice-Hall.

Jonassen, D. H., Carr, C., & Yueh, H. P. (1998). Computers as mindtools for engaging learners in critical thinking. *Tech Trends, 43*(2), 24-32.

Kirschner, P. A. (2000). Using integrated electronic learning environments for collaborative teaching/learning. *Research Dialogue in Learning and Instruction, 2*(1), 1-10.

Kirschner, P. A., & Selinger, M. (2003). The state of affairs of teacher education with respect to ICT. *Technology, Pedagogy and Education, 12*(1)*,* 5-17.

Kirschner, P. A., & Davis, N. (2003). The Pedagogic benchmarks for ICT teacher education. *Technology, Pedagogy and Education, 12*(1), 127-149.

U.S. Congress, Office of Technology Assessment. (1995). *Teachers and technology: Making the connection* (OTA-HER-616). Washington, DC: U.S. Government Printing Office.

Pearson, J. & Selinger, M. (1999). Linking different types of knowledge in professional education and training: the potential of electronic communication. In M. Selinger and J. Pearson (Eds.), *Telematics in education: trends and issues* (pp. 15-31). Oxford, England: Elsevier.

Preparing Tomorrow's Teachers to Use Technology (PT3). U. S. Department of Education. Retrieved on May 2, 2002 from http://www.ed.gov/teachtech/

Reeves, T. C. (1994) Evaluating what really matters in computer-based education. Retrieved on May 2, 2002 from http://www.educationau.edu.au/archives/cp/reeves.htm

Reeves, T.C. (2000). Alternative assessment approaches for online learning environments in higher education. *Journal of Educational Computing Research, 23*(1), 101-111.

Selinger. M. (2001). How can ICT increase the recruitment, retention and morale of teachers? In M. Selinger and C. Yapp. *ICTeachers.* London: IPPR.

UNESCO (2002). *Information and Communication Technologies in teacher education: A planning guide*. Paris: UNESCO.

Yapp, C. (2001). Reinventing the Teacher: The impact of pervasive technology. In M. Selinger and C. Yapp. *ICTeachers.* London: IPPR

Appendix 1. Web site with more information on the 26 cases of good ICT in teacher education.

	Teacher Education Program	Country	URL(s)
1.	Bachelor of Education (Primary), University of Western Sidney	Australia	http://www.uws.edu.au/courses/strand_course?COURSE=1
2.	Bachelor of Education and Teaching, University of Tasmania	Australia	http://www.educ.utas.edu.au/users/afluck/IT/PersProf
3.	Curtin University of Technology, Perth, Western Australia	Australia	http://www.edna.edu.au/sibling/tefa/html/body_attachment_
4.	Galileo Educational Network	Canada	http://www.galileo.org
5.	eScience (also known as ASTeL*)	Finland	http://www.malux.edu.helsinki.fi/malu/koulutus/escience/in
6.	Sipoo Institute	Finland	http://sipooinstituutti.net
7.	Technology Oriented Teacher Education Programme	Finland	http://wwwedu.oulu.fi/tekno/Eesittely.html
8.	The Finnish Virtual School for Science Education (FVSSE)	Finland	http://www.malux.edu.helsinki.fi/malu/tutkimus/index.en.ht
9.	La main à la pâte	France	http://www.inrp.fr/lamap -in French
10.	eL3 eLearning and eTeaching in Initial and Further Teacher Education	Germany	www.eL3.de (in German)
11.	Learning and teaching on the Internet, Israel	Iceland	http://www.asta.is/nkn2
12.	The teaching of children's literature	Iceland	http://barnung.khi.is
13.	Learning and teaching in the 21st century	Israel	No website available
14.	Networks for the teaching methodology of foreign languages	Netherlands	http://www.ilo.uva.nl/english
15.	The New Style Teacher Training Program for Elementary Education	Netherlands	http://www.nldata.nl/ict-e/English/ c
16.	BA(QTS), Bachelor degrees with Qualified Teacher Status	UK	http://www.warwick.ac.uk/wie/spo/index.htm
17.	Cambridge Consortium ICT Training for Teachers	UK	http://www.cea.co.uk/ict.htm - Not public domain
18.	Post Graduate Certificate of Education	UK	http://www.socsci.ulst.ac.uk./education/
19.	Postgraduate Certificate in Education Primary course, University of Brighton	UK	Not public domain
20.	Strategic Leadership of ICT	UK	http://www.ncsl.org.uk
21.	The Learning Network	UK	http://www.learning-network.net/
22.	ASU West Early Childhood Teacher Preparation Program	USA	http://westcgi.west.asu.edu/pt3/publications/NETSTsc.htm
23.	ISU Elementary Education	USA	http://www.educ2.iastate.edu/ci/classes/201/
24.	ISU Graduate Education in Curriculum & Instructional Technology	USA	http://www.ctlt.iastate.edu/
25.	University of Texas UTeach program	USA	http://www.uteach.utexas.edu/technology/
26.	University of Virginia Secondary Education programs	USA	http://www.ncate.org/accred/projects/tech/tech-21.htm

Authentic Learning in the Schools: Teacher Practices, Attitudes, and Challenges

Theodore J. Kopcha
Arizona State University

Introduction

The idea of authentic learning has received considerable attention since its popularization by Brown, Collins, and Duguid (1989). Drawing from Lave (1988), they argue that learning is a process of enculturation, where what you learn should not be separated from the context or situation in which it is normally used. Being submersed in activities that pertain to this context allows the learner access to the social structure that surrounds it, and "these coherent, meaningful, and purposeful activities are authentic" (Brown, Collins & Duguid, 1989, p.34).

Authentic learning, then, is a product of situated cognition and authentic activity. Performing authentic activities provides the learner with contextual cues that typical classroom learning often lacks (Brown, Collins & Duguid, 1989). These cues help the learner encode knowledge in conjunction with its purpose, increasing the learners' likelihood to draw upon and use such knowledge to solve future problem situations (Brown, Collins & Duguid, 1989; Hannafin & Choi, 1995; Cognition and Technology Group at Vanderbilt, 1990; Herrington & Oliver, 2000). As a result, the focus of learning turns to the process rather than the outcome (Hannafin & Choi, 1995). Several approaches to authentic learning have emerged, including cognitive apprenticeships (Brown, Collins, & Duguid, 1989) and Herrington and Oliver's (2000) instructional design framework for authentic learning environments. Authentic multimedia environments such as the Vanderbilt group's Jasper videodisc series (1990) have attempted to bridge the gap between theory and practice as well.

Despite the inherent attraction of authentic learning, the research surrounding the benefits associated with teaching in an authentic manner is limited. Much of the research is restricted to more naturalistic studies and is often contradictory. In a study of eight pre-service students, Herrington and Oliver (2000) found that, for complex learning, a multimedia situated learning environment could be just as effective as more traditional teaching approaches. By combining authentic contexts, activities, and assessments into a multimedia environment, they found that students felt appropriately challenged and had a clear understanding of how the theory they were studying tied into the activities they performed, though a majority of students felt the time allocated to do so was insufficient.

Welhage, Newmann, and Secada (1996) define authentic learning by separating it into three parts: construction of knowledge, disciplined inquiry, and value beyond school. They suggest that combining authentic contexts, activities, and assessments can improve student achievement when used properly. However, in a study of 10 gifted high school students, Bishop (2000) observed a teacher's use of these three measures as a guide to teaching and evaluating students who performed independent research projects. She found that, when applying these measures, only three of the students were able to meet all three requirements of authentic learning. In-depth understanding and the expression of conclusions were most often missing from student work, and most of the explanations lacked the depth expected from teaching in such a manner. Most importantly, she noted that these students may have had a more authentic experience if they had been given initial guidance regarding the process of independent research. This finding is in line with the concerns of Anderson, Reder, and Simon (1996), who review evidence indicating that the gains provided by teaching in a situated manner may be achieved other ways.

This study of authentic learning in the schools was one key component of a three-year, million-dollar project that analyzed the use of desired instructional planning and practices, including authentic learning, in 11 elementary schools across the state of Arizona. The evaluation of each school program included the investigation of teacher use of authentic learning activities and attitudes toward these activities.

Method

Participants

The participating schools were chosen by staff members of the Arizona Community Foundation, which funded the project, and by consultants to the foundation. Schools were primarily of lower and middle socioeconomic status and were selected to represent the various geographic areas across the state. The schools were participants in a learning communities project with the goal of involving parents, teachers, and the community in school decision-making.

The 11 participating schools were spread throughout the state, spanning nine different school districts. Each school was selected for funding based on an extensive proposal that included a plan for achieving the goals of the project. Acceptance into the project included a stipend of $35,000 per year for each school over the three-year period of the project.

A total of 47 teachers, administrators, and parents were interviewed regarding authentic learning during the final year of the project: 29 were site team members consisting of teachers, administrators, and parents, and 18 were teachers who did not participate on a site team. The site team members participated in regular meetings to make sure the school was meeting its goals for the project, as well as to address other factors related to the goals. The team members were volunteers and each team included at least two members of each type. Both teachers and administrators ranged widely in levels of work experience.

Procedures

As part of a formal evaluation performed during the last year of a three-year project, four doctoral students in Educational Technology at a major university in the southwest acted as experimenters and collected data about the school practices and activities using a variety of methods. These included two one-day site visits to each of the 11 schools for the purpose of observation and information gathering and between eight and ten interviews per school with school principals, site team members, teachers, and parents using a structured interview protocol. One focus of the school visits and interviews was to identify teacher use of authentic learning activities, their attitudes toward authentic learning and their perceived challenges with respect to using authentic learning activities in the classroom.

The evaluators conducted individual interviews with each of the 18 teachers and the 29 site council members, many of whom also were teachers. They used the standard interview protocol for teachers or site council members as appropriate, conducting approximately half the interviews on a face-to-face basis and the other half by telephone. One purpose for the site visits was to collect documents and observe projects or activities that related to authentic learning and other project activities. The site visits and interviews were conducted during the fall of 2002, the last semester of the three-year project. The site visits were conducted once at the beginning of the fall semester, and once at the end of the semester.

Criterion Measures

Data were collected using five instruments: a School Report Form; separate, individualized interview protocol for site team members, teachers, and parents; and a site observation form. The teacher and site council member forms consisted primarily of open-ended questions and the site observation form instructed the evaluators to look for specific evidence of authentic learning practices. The content was created based on the goals of the project and through extensive feedback with the funding agency.

The teacher interview protocol consisted of 18 items and the site team member interview protocol consisted of 24 items. The questions asked how authentic learning was incorporated into instruction, the effect it has had on student achievement, the training received and interest in additional training regarding authentic learning, as well as asking for examples of authentic learning in the teachers classrooms and the school's success in doing so on a curricular level.

Once the data were collected, the evaluators created general statements based on recurring themes, paraphrased the original statements to broaden the category, and tabulated the number of responses that fell into a particular category. When a response was unique, it was written as it was expressed in the interview. However, when multiple unique responses were given, the evaluators would report a representative number of those responses, to convey the variety of responses without including every one. For reported authentic activities, the data were separated by high and low levels of authenticity, based on the judgment of the four evaluators.

Results

As shown in Table 1, teachers responded 'yes' with the highest frequency, 17 out of 18, to the following questions: "Do you incorporate authentic learning tasks and/or assessment into your instruction?" and "Do you feel your use of authentic learning tasks and/or assessment has affected student achievement?". The second most frequent 'yes', 16 out of 18, was generated from the question "Would you like additional training regarding authentic learning and assessment". The least frequent 'yes', 11 out of 18, pertained to prior training in authentic learning, where many had been exposed to authentic learning in a college curriculum in education or in professional development.

Table 1 *Teacher and Site Council Member Responses to Authentic Learning Items*

Interview Items	Yes	No	Sample Response
Teachers (N=18)			
Do you incorporate authentic learning tasks and/or assessment into your instruction?	17	1	• Data collected and analyzed for use in multi-disciplinary tasks
Do you feel your use of authentic learning tasks and/or assessment has affected student achievement?	17	1	• Deepens the students' understanding • Increases motivation to learn
Have you received training for or assistance with incorporating authentic learning tasks and assessment into your instruction?	11	7	• Professional development • Training received in college
Would you like additional training regarding authentic learning and assessment?	16	2	• Explain authentic learning • Provide examples and models
Site Team Members (N=29)			
Do you feel the school has been successful in incorporating authentic learning tasks into the curriculum?	24	5	• More emphasized not but always did this • Projects are more authentic now

Two of the teachers cited collecting data and analyzing it for use in multi-disciplinary tasks as an example of incorporating authentic learning tasks. When asked about the effects of authentic learning on student achievement, six teachers included deepening student understanding and six included increasing motivation to learn. When asked about the additional training and what it should encompass, eight of the responses included a combination of explaining what authentic learning is and demonstrating model lessons or examples of authentic learning.

Site team members were asked if authentic learning was successfully incorporated into the curriculum. The majority of members, 24 out of 29, responded 'yes', stating that the projects they observed teachers and students completing were more authentic than before the project.

Table 2 displays all of the teachers' reported authentic activities, organized by level of authenticity. Only five of the tasks reported by the teachers were judged as high in authenticity by the evaluators. At two different sites, one teacher reported that students built a garden, planted and grew seeds for the garden, and maintained the garden to use data collected from the garden in science and math. One of the two teachers regularly took students off site to collect data from natural settings, such as nearby National Parks. The data were then used to create graphs that mathematically represented the data collected, as well written reports that combined the skills learned in science and language arts. The evaluators regarded these activities as being high in authenticity.

Table 2 *Teacher's Reported Authentic Learning Activities by Level Of Authenticity* (N=18)*

Level of Authenticity	Number of Instances	Description of Activity
High	2	• Students participate in the creation of a garden and share the responsibility of maintaining that garden as part of a unit on habitats.
	2	• Bus trips taken off school grounds to collect data. The data were then used in science, math, and writing.
	1	• Students speak to a video camera on social skill topic and show the video to school.
	1	• A student operated postal system for the school.
Low to non-authentic	3	• Discuss how a particular lesson relates to everyday life.
	2	• The math program we use "Math In Context". This uses authentic activities.
	1	• When learning to tell time, I stop and ask students what time it is.
	1	• Reports written with a focus, such as penguins.
	1	• General group work.
	1	• Integrate students with special needs in the classroom.
	1	• Use manipulatives to learn math, incorporate math with science project, and writing in science journals.
	1	• Align assessment to instruction.
	1	• Student role playing during Martin Luther King lesson.
	1	• Study an artist, go to the art museum, then do artwork back at school.

*Level of authenticity was a judgment made by the evaluators.

A popular activity reported by teachers as authentic was telling students how the content they were learning fit into their own lives. As shown in Table 2, the evaluators rated this activity as low in authenticity or non-authentic. Other selected examples of activities rated low in authenticity included using manipulatives in mathematics, stopping class to ask what the time is, integrating students with special needs into the classroom, and aligning assessment with instruction.

Discussion

The most important finding of this study pertains to teacher's reported use of authentic activity. While the examples of tasks collected by the evaluators varied greatly in nature and content, very few of the tasks could be judged as solid examples of authentic learning tasks. In fact, only five of the 18 teacher-generated examples of authentic learning tasks were considered highly authentic by the evaluators, while the remaining 13 were considered low in authenticity or non-authentic.

There are several possible reasons for the low number of authentic learning tasks. One is that many of the teachers interviewed were unfamiliar with what an authentic learning task is or what authentic learning is. Slightly more than half the teachers (11 out of 18) had received formal training in authentic learning, either through a college preparation program or a professional development setting. Since authentic learning has become popular only during the last 10-15 years, it seems likely that many of the teachers interviewed may not have received an opportunity for training regarding authentic learning and therefore may not have been familiar with the concept. Teachers, then, may have been reporting activities that they truly felt were authentic based on their personal experiences and training in authentic learning, which was limited for many of them.

A second reason for the low number of authentic learning tasks may be that more recent changes in education have increased the demand for standardized test scores as a measure of growth and accountability for both teachers and students. Teachers commented on having difficulty with reconciling demands for incorporating authentic learning tasks with the school concerns about achievement on standardized or state-mandated tests. In *Successful School Restructuring*, Newmann and Welhage (1995) noted this, where restructuring schools for authentic teaching often fails due to concerns about other school goals.

A third possible reason may be that the schools in the evaluation were all elementary schools. Much of the content that students study in elementary school is basic or foundational. This may have made it difficult for teachers, as well as other adults, to conceptualize the rather complex tasks that are "authentic", since such tasks are often the tasks that adults and older youths perform regularly in the real world. Tasks such as reading books, writing letters and some types of reports, and solving realistic math problems seem authentic for elementary school children, yet are not considered authentic by some advocates of authentic learning. Often, these age- and level-appropriate tasks do not come to teachers' minds when they are asked to report authentic learning tasks since they do not meet the level of complexity that makes the activities of adults and older youths appear authentic.

These reasons may help explain why many of the examples of authentic learning tasks are small, short connections with the real world, or small additions to more traditional classroom activity. A teacher faced with the simultaneous demands of preparing students for a standardized test and of teaching in an authentic manner might try to combine the two. A combination of those might look like the examples judged as 'low' in authenticity, such as creating a report with a focus or simply discussing how a particular lesson relates to everyday life.

Another important finding of the study was the highly positive response teachers and site team members displayed regarding the incorporation of authentic learning tasks in the curriculum and classroom. The majority of site team members (24 out of 29) felt they had been successful incorporating authentic learning tasks into the curriculum, and nearly all the teachers (17 out of 18) reported that they had incorporated authentic learning tasks into their instruction and that doing so had improved student achievement. However, when the evaluators judged the reported examples of authentic learning tasks, few of them were highly authentic.

As mentioned previously, the fact that the teachers felt the tasks were indeed authentic may be based on their limited understanding of authentic tasks. However, enthusiasm regarding the incorporation of authentic learning tasks also may be due to the nature of the data collected. The information collected was self-reported data, and data collected in this manner is subject to self-presentation bias (Igoe & Sullivan, 1993; Shen, Sullivan, & Igoe, 1996). Both teachers and site team members, then, may have eagerly answered 'yes' when asked if they were incorporating authentic learning tasks in order to appear that they had fulfilled the obligations of the grant. When pursued further, however, they often were unable to generate sound examples of authentic tasks according to the evaluators' standards. This also provides some explanation for the first finding, where many of the reported authentic learning tasks were judged as low in authenticity.

The examples judged as 'high' in authenticity shared two common elements that the others lacked: appropriate amounts of time and resources. The two teachers who had students building gardens and using those gardens for data collection and analysis devoted ample time to participating in such tasks. At one particular site with a small student population, teachers would collaborate on a weekly basis to create or locate existing authentic learning activities and monitor students' progress throughout the year. That time was used to generate goals for each student and assess progress towards the previous goals. Newmann and Welhage (1995) support this, suggesting that successful restructuring towards authentic teaching is easier with small school size, as well as opportunities for teachers to collaborate and continually assess their progress towards teaching in an authentic manner.

Financial resources are necessary to engage in authentic learning activities as well. For the examples that were 'high' in authenticity, a portion of their annual stipend of $35,000 was spent on materials, transportation, teacher training, and supplies necessary to perform authentic tasks. The sites that participated in the highly authentic learning tasks provided professional development on authentic learning funded by outside sources. The activities such as building and maintaining a garden or taking a group of students to a nearby national park were paid for by external funding sources as well. The majority of teachers in the study, though, were not provided with such resources.

Finally, the same number of teachers who responded positively to incorporating authentic learning tasks and to its effect on student achievement requested additional training in authentic learning.

Specifically, they asked for an explanation of authentic learning and more examples or model lessons of authentic learning tasks. This further supports the idea that teachers may have been generating examples of authentic learning tasks based on a limited understanding of the concept and that little effort was spent on training most teachers about authentic learning tasks.

As evaluators of the authenticity of teaching practices, the evaluators also encountered difficulty in judging the authenticity of the tasks. Despite many definitions and "examples" in the literature, few of the reported examples seemed authentic to us. The examples from this study often contained some elements of authentic learning but not others, stressing the limits of the definitions that exist. The evaluators in the group often debated between two philosophical extremes: perhaps most or all learning is authentic or perhaps very little learning by young children meets the standards of some advocates of this approach. It seemed to us that the term "authentic learning" (Brown, Collins & Duguid, 1989; Hannafin & Choi, 1995; Lave, 1988) is used too much in the literature to deride the practices of well-intentioned, good teachers.

Contributing to this issue is the lack of a well-articulated, pragmatic teaching approach for authentic learning. While Newmann, Secada, and Welhage (1996) provide three standards for authentic teaching practices, the authentic learning tasks and assessments they describe often contain some of the key features of authentic learning while lacking others. The current approaches to situated learning increase this difficulty since they often involve a level of complexity that is not authentic or that the average teacher is unable to achieve, especially without training and considerable resources.

Conclusion

Overall, the teachers and site team members involved in the project were consistently described as hard working and as doing a good job of educating their students. A tremendous amount of support and assistance is required to accomplish a transition to teaching in such a manner as this, and many teachers did the best they could to meet the goals of the project while meeting the demand to improve student achievement. While the everyday work of the schools often took priority over implementing authentic learning tasks and assessments into the curriculum, it was consistent with the goal of good student achievement.

References

Anderson, J.R, Reder, L.M., & Simon, H.A. (1996). Situated learning and education. *Educational Researcher,* 25(4), 5-11.

Bishop, K. (2000). The research processes of gifted students: a case study. *Gifted Child Quarterly, 44(1)*, 54-64.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *EducationalResearcher, 18(1),* 32-42.

Choi, J., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research and Development*, *43(2),* 53-69.

Cognition and Technology Group at Vanderbilt (1990). Anchored instruction and its relationship to situated cognition. *Educational Reseacher, 19(6)*, 1-10.

Herrington, J. & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, *48(3),* 23-48.

Igoe, A. & Sullivan, H. (1993). Self-presentation bias and continuing motivation among adolescents. *Journal of Educational Research, 87(1),* 18-22.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life.* NY: Cambridge University Press.

Newmann, F.M., & Welhage, G.G. (1995). *Successful School Restructuring: A report to the public and educators by the Center on Organization and Restructuring of Schools.*. Madison, WI: Center on Organization and Restructuring of Schools.

Newmann, F.M., Secada, W.G., & Welhage, G.G. (1996). *A guide to authentic instruction and assessment: Vision, standards, and scoring*. Madison, WI: Center on Organization and Restructuring of Schools.

Newmann, F.M., Secada, W.G., & Welhage, G.G. (1996). Standards for authentic achievement and pedagogy. In Newmann, F.M., & Associates, *Authentic Acheivement* (pp.21-48). San Francisco: Josey-Bass Publishers.

Shen, S., Igoe, A., & Sullivan, H. (1996). Self-presentation bias and continuing motivation among Chinese students: a cross-cultural phenomenon. *Journal of Educational Research, 90(1),* 52-56.

Heuristic Task Analysis on Expertise in Designing Web-Based Instruction

Ji-Yeon Lee
University of South Carolina

Charles M. Reigeluth
Indiana University

Abstract

To better understand tacit knowledge underlying experienced Web-Based Instruction designers' performance, this study utilized Heuristic Task Analysis, a method developed for eliciting, analyzing, and representing expertise in complex cognitive tasks. Three experts, representing three post-secondary institutions offering online programs, were selected for interviews based on the following criteria: 1) WBI design knowledge and skills indicated by formal training/education and experience (i.e., years of experience and number of courses designed); 2) peer recognition; and 3) availability. The interviews focused on three different WBI design cases (one from each institution), and the experts were asked to articulate their underlying thoughts and principles for designing online courses. On the surface, the WBI design processes that the experts used looked more alike than different, entailing major steps such as meeting with faculty, developing content, monitoring courses in progress, and debriefing the instructor and students. The underlying principles and knowledge that guided each expert through the processes, however, were unique in that each expert constructed her own heuristics to accommodate the myriad contextual factors that arose in her work setting (e.g., student population, program size and structure, expert backgrounds).

As the number of higher education institutions offering online courses rapidly increases, design and development of Web-based Instruction that is pedagogically sound and cost-effective is of special concern for many researchers and practitioners in postsecondary institutions. On the surface, because WBI has become so prevalent in higher education and the advantages of Web-based learning seem so powerful, proponents of WBI assume that we know how to design and develop successful WBI and that we have an accurate understanding of the processes we seek to support and enhance. As the history of instructional media shows us, however, technological potentials do not easily transfer into direct educational benefits (Cuban, 1988; Zhao & Parks, 1995). Indeed, designing and developing WBI that lives up to the standards of traditional face-to-face instruction is certainly more complex and much less hopeful than a surface view would reveal (Carr-Chellman & Duchastel, 2000). Thus, the purpose of this study is to explicate the expertise underlying the expert performance of designing WBI, and then based on the expertise, to develop practical guidelines to facilitate design and development of quality WBI.

Heuristic Task Analysis

Heuristics, mostly known as rules of thumb, educated guesses, or intuitive judgments, are strategies and criteria for deciding the most effective courses of action among several alternatives in problem solving situations (Pearl, 1984), such as designing Web-based instruction. Thus, *heuristic tasks* are "tasks for which experts use causal models—interrelated sets of principles and/or guidelines—to decide what to do when, such as a high school course on thinking skills or a corporate training program on management skills" (Reigeluth, 1999, p. 435). As our society in general and the workplace in particular become more complex, our abilities to deal with heuristic tasks become more critical than ever before. Thus, powerful methods for analyzing the often tacit knowledge underlying complex, heuristic tasks are in great need.

Heuristic Task Analysis (HTA) is a method developed for eliciting, analyzing, and representing expertise in complex cognitive (or heuristic) tasks (Reigeluth, 1999). Traditional task analysis methods typically result in a list of tasks and subtasks stated in behavioral terms and framed in a procedural format (Gordon & Gill, 1997; Ryder & Redding, 1993). In contrast, the HTA method primarily focuses on analyzing relatively complicated, ill-structured tasks for which experts use such heuristic knowledge as principles, causal models, and guidelines to decide what to do when, instead of being aware of and deliberately following a set of steps.

The HTA is conducted by asking the question, "What is the simplest version of the task that an expert has ever performed?" and "What is the next simplest version?" and so on. Identifying the boundary of the simplest yet still representative version of the task dramatically reduces the amount of knowledge and number of contextual factors that are necessary for expert performance within that boundary. More importantly, starting with the simplest version allows analysts to study the phenomenon of expert performance under simpler conditions and to do so in a reliable fashion (Ericsson & Charness, 1997). Once the simplest version is analyzed, then progressively more complex versions of the task can be determined and analyzed in order to identify the additional knowledge that experts use to perform them. (For more detail, see Appendix A).

Methods

The task selected for analysis in this study is the design of WBI for distance students in university settings. WBI design has some procedural but mostly heuristic elements, and this study identifies both, with particular focus on the latter because they have been found to be more pervasive in this complicated process that involves numerous variables and design considerations (Julian, Larsen, & Kinzie, 1999; Lohr, 1998).

Selection of Experts

Previous studies addressing the issue of amassing sufficient numbers of experts for knowledge acquisition advise that the scope and depth of interviews eventually depends on the purpose and practicality of any given study (Shadbolt, O'Hara, & Crow, 1999). Since the purpose of this study was to identify in-depth heuristics of experts in designing WBI, rather than to generalize the heuristics, we interviewed fewer experts, just enough to avoid idiosyncrasy, and interviewed each expert in more depth until the saturation point was reached.

For conducting the HTA (see Appendix A), we identified three experts based on three criteria – relevant knowledge and skills, peer recognition, and availability. Table 1 summarizes the qualifications of the three participating experts.

Table 1. *Participating Expert Profiles*

Name	*Title*	*Highest Degree Earned*	*Experience*	*No. of Courses*
Catheri ne	Instructional Designer	Masters in Instructional Technology	3 years	16
Michelle	Instructional Designer	Masters in Instructional Design (Ph.D. in progress)	4 years	125
Anita	Instructional Designer	Masters in Instructional Design	3 years	40

Note: Pseudonyms were used for all participating experts.

Knowledge and Skills

The first selection criterion was whether the person had the actual knowledge and skills that were needed to design and develop successful WBI courses. The knowledge and skills could have been gained from two sources – formal training and experience.

Given the lack of knowledge about how to identify heuristic expertise, having degrees or certifications in instructional design or other related fields does not indicate that one has the necessary knowledge and skills to design good WBI. In fact, some of those who have designed and developed very successful WBI courses may not have received much formal training. However, when interviewing experts to elicit their tacit knowledge and skills, it is important that they have the vocabulary to verbally represent their knowledge and skills. We argue that formal instructional design training helps people to become more consciously aware of the knowledge and skills underlying their WBI design and to have a better vocabulary to describe the knowledge and skills. Thus, we selected those experts who had advanced degrees or training in instructional design or related fields for interviews.

The experience of experts was determined by the number of years they had designed WBI courses and the number of WBI courses they had designed. We selected experts who had designed and developed at least 16 WBI courses in university settings during more than 2 years. Simon and Chase's classic study on expertise argues that a 10-year period of intense preparation is necessary to reach the level of "expertise" in most domains. This figure of 10 years has been accepted by many researchers as a criterion to determine one's expertise (Ericsson & Charness, 1997; Simon & Chase, 1973). However, given the fact that the World Wide Web has only become generally available to the public since about 1996, this so-called "10-year rule" was not applicable in our case, and we modified the criterion accordingly.

Peer Recognition

To elicit heuristic knowledge for designing successful WBI, it is crucial to select experts who have successfully designed and developed such courses. One way to distinguish the most successful WBI courses is soliciting recommendations for outstanding courses from peer instructional designers or instructors. Thus, at professional conferences on instructional design, distance education, or related topics, the first author asked professional instructional designers and faculty who designed and developed WBI courses in university settings to recommend peers they believed to be expert WBI designers. Once a number of experts were recommended, the first author made contact and asked for access to the best courses that each had developed. As more and more universities and colleges were using course management systems (e.g., WebCT, Blackboard) with password-protected access, we were unable to access course materials before in-person interviews with experts. However, all participating experts provided some brief information about their exemplary WBI courses before the interviews.

Availability

The HTA process is a very demanding and time-consuming process for both the analyst and the participating experts. Thus, the last but not least criterion was the availability and willingness of the experts to participate in multiple rounds of in-depth interviews and exert their mental energy to explicitly elaborate the tacit knowledge and skills underlying their performance of designing and developing WBI.

Data Collection Methods

Interviews

Like most knowledge elicitation studies involving human experts as data sources, in this study, semi-structured, in-person interviews were used as the primary data collection method.

Studies of experts emphasize the importance of the use of in-person, on-site interviews in order to uncover their heuristic knowledge, because experts have been found to better recall details of their performance when they are actually in their work environments (Benbenishty, 1992; Hoffman, Shadbolt, Burton, & Klein, 1995; Kitto & Boose, 1987; Olson & Rueter, 1987; Yi, 1992). Thus, the first author visited each expert in his/her office and conducted interviews, which were audio taped and transcribed for analysis. When it was a phone interview, she called the experts at their office when it was most convenient for them.

Each expert was interviewed a minimum of three and maximum of five times, for about 60-90 minutes each time. It was estimated that a total of 30 hours would be necessary to elicit the in-depth heuristics of all of the experts. This was a tentative, approximate number of hours that was based on an estimate of expected reasonable depth of coverage of the expertise given the purpose of the study. The number of interviews for each expert was adjusted in the course of the investigation (Merriam, 1998). In other words, the interviews were continued until a point of saturation or redundancy was reached.

Once experts agreed to participate in the study, and before in-person interviews, we asked them to complete an open-ended e-mail survey. The purpose of this survey was to collect background information about each expert to prepare for in-person interview protocols. The survey had 11 questions in three categories – WBI design experience, work environments, and WBI courses they had designed.

The interviews with each expert centered on eliciting information about the simplest version of the task that s/he had designed and delivered. This allowed us to understand the case-based reasoning that was used by the expert to identify and represent his/her expertise (Schank, 1982; 1989). In previous case studies, we found that using a specific case as a basis of discussion was instrumental to explicating tacit knowledge through using a well orchestrated array of observations, descriptions, and probing questions (Lee & Reigeluth, in press).

The HTA offers a set of interview questions for probing the expert's heuristic knowledge for performing the task of designing WBI. We were not restricted to the predefined questions, and the overall interview process was flexible, emergent, and reflective in nature. In other words, depending on the strengths and problems of the initial interviews in identifying task expertise, we revised the HTA method for each subsequent interview. The HTA method is an iterative process, which means that finishing the first round of HTA is not the end of the study but the beginning of the second round of HTA; and the end of the second round is, again, the beginning of the third round; and so on.

Data Analysis and Interpretation

Triangulation. For more thorough data and greater breadth, this study involved multiple experts as data sources. The task of WBI design is often perceived as collaborative teamwork requiring more than one person (i.e., instructional designer) for completion. When comparing findings from the interviews with each expert, we encountered situations in which different experts approached the same problem with totally different solutions that were equally effective. In those cases, rather than trying to seek a single best solution, we tried to identify the different conditions under which each solution worked well.

Member checks. After each interview with an expert, the first author transcribed the interview and brought a summary of the interview and tentative interpretations to the next interview for review prior to the next interview. Through this process, the experts corrected the researcher's errors and misconceptions about their expertise, and she had a chance to ask additional questions to clarify the data already collected.

Three WBI Design Cases

This section reports findings from three WBI design cases. Each case begins with a description of background information (i.e., program/course setting and expert biography); next, it explores procedural and heuristic knowledge identified by the expert; and it concludes with a discussion of case-specific findings.

Case 1: E-MBA Program at Purdue University

The e-MBA program at the Krannert School of Business, Purdue University, was a 2-year old, fully accredited, online executive MBA program in Food and Agricultural Business. As an alternative to traditional part-time executive MBA programs, the e-MBA program heavily utilized Web-based distance learning technology allowing full-time working professionals to pursue degrees. The program consisted of four semester modules, 22-weeks each, and students progressed in a cohort of 25 by taking three courses per semester. Although the majority of faculty-student and student-student interactions occurred online, the program did have residency components. Students were required to be on campus for 9 weeks during their program, including a 1-week, on-site orientation and 2 weeks every semester.

Catherine was an instructional designer at Purdue. Since she had accepted the position of Distance Education Specialist for the e-MBA program in early 2000, she had guided faculty through the process of transforming face-to-face courses into Web-based courses, managed course delivery for two ongoing student cohorts, and provided technical training. She was in charge of every course in the program and had designed and revised 16 online courses.

The course Catherine identified as the simplest yet most representative of her WBI design work was "Applied Quantitative Analysis." This course explored the application of contemporary concepts and quantitative techniques for decision making in the context of food and agricultural business management. Applied Quantitative Analysis had been taught for 2 years by the same instructor with 16 and 27 students respectively, and was initially developed by another instructional designer. There were numerous student complaints about the original course, mainly due to its complex, unorganized interface and unrealistic student workload. Thus, Catherine worked with the faculty to restructure the existing WBI course with a more intuitive interface and redesigned student assignments.

Case 2: Continuing Education and Special Programs at the University of Northern Iowa

Since 1996, the University of Northern Iowa (UNI) offered online courses through the Continuing Education and Special Programs (CESP) office to expand access to educational offerings for students who lived beyond commuting distance and wished to pursue degrees or update their knowledge in specific content areas. Students who took Web courses through the CESP were mostly in-service teachers working on their master's degrees. As a cohort, they usually registered for one course per semester, and it took them up to 3 years to complete the program.

Michelle was an instructional developer with the CESP at the UNI. She had designed and developed numerous WBI courses since 1997: she began as a graduate student, assisting faculty with online course development, and later accepted a staff position and became a full-time instructional designer.

The course Michelle identified as the simplest yet most representative of her WBI design work was "Communication Theory in Media." Exploring contemporary theories of mass communication, learning, perception, and propaganda as they applied to message design utilizing communication media, this course was an elective in the master's program in Educational Technology and a required course in the Master's program in Communication and Training Technologies. There were approximately 20 students enrolled in the course during the Spring 2002 semester. About half of the students were teachers at various locations in the state of Iowa working on their master's degrees online, and the other half were residential graduate students.

In previous semesters, this course had been only taught face-to-face and always by the same instructor. Like most WBI courses she designed, Michelle was given two months to work with the instructor and coordinate the design and delivery of the course. The graphic designer who recently had joined the CESP provided assistance with graphics (i.e., course banner and customized icons for the course website), but Michelle did most of the work, including consulting with the faculty member, converting existing course materials, designing the course website, and training the faculty.

Case 3: Educational Outreach at the University of South Florida

The vast majority of distance learning initiatives at the University of South Florida (USF) grew rapidly in terms of both student enrollments and technology integration since 1998, and most online courses were coordinated through the division of Educational Outreach. Anita joined Educational Outreach as a graduate assistant in 2000. Since then, she had provided course support for faculty teaching online courses and guided the design and delivery of numerous courses at USF. As a professional instructional designer, her major responsibilities included meeting with faculty to discuss design issues and on-line options for their courses, acting as liaison between faculty and the USF library to secure copyright permission for content, helping faculty prepare and upload materials to course websites, and maintaining/updating/trouble-shooting websites.

The course Anita identified as the simplest yet most representative of her WBI design work was "General Chemistry (I)." This course explored principles and applications of chemistry such as properties of substances and reactions, thermo-chemistry, atomic-molecular structure and bonding, and periodic properties of elements/compounds, and it was a multiple-section course taught by six different instructors. Since it was a required course for every first-year student in the chemistry department, more than 1,000 students (200 students for each section) were enrolled during the Fall 2000 semester.

General Chemistry (I) was problematic to convert online because of its volume of students; there had been some improvement in terms of WebCT server capability, but at the time Anita was working with the instructors, the server could not handle 1,000 students accessing a course shell at the same time. Thus, the course was structured utilizing a main course shell, which served as an umbrella website for the entire class and was linked to separate course shells for each section.

Comparison of Findings across the Three Design Cases

Factors that Affected the WBI Design Process

When asked to identify major factors that affected task difficulty in designing WBI courses, all three experts responded that faculty, students, and content (i.e., types and amount of course materials) were the most important considerations.

Faculty. The experts were unanimous in confirming that faculty was probably the most crucial factor to consider when designing online courses. The underlying reasons why each expert considered faculty important in WBI design, however, differed depending on her context. Catherine was concerned about the mindset of individual faculty. She put considerable effort into creating a consistent design across the curriculum and disseminating best practices of online pedagogy. Thus, individual faculty members' openness to take suggestions and adopt previously established program standards was important for her role as the program instructional designer. For Michelle, given the large volume of courses assigned to her each semester, the timeliness and cooperation of faculty to provide course materials and feedback became a critical factor. Anita took a role as consultant when working with faculty. Her priorities were meeting individual faculty members' needs and desires as well as those of students, so she considered faculty members' ability to conceptualize their courses and communicate their needs most important.

Students. It is not surprising that all three experts identified students as another crucial factor in determining their approaches to WBI design. Whether face-to-face or online, most decisions instructional designers make are informed by who their audience is, and the three cases studied were no exceptions. Besides student prerequisites and individual learning styles that are usually considered important in the learner analysis literature, student experience with the particular programs that they were enrolled in and with other online courses, as well as their technical proficiency and access to equipment/bandwidth, became factors for WBI course design. Especially when students were taking courses as a cohort, as in the cases of Catherine and Michelle, designing courses for second- or third-year cohorts became much easier because they could assume the level of student knowledge based on courses students had (or had not) taken.

Content (Types/Amount of Materials). For Catherine and Anita, content was considered an important factor because, depending on the types or amount of materials to be converted or created for the Web, the development/production time and effort could vary from a matter of 1-2 day(s) to up to 2-3 months. Michelle was more concerned with how much time she needed to develop a new course; but considering that the types and amount of materials (e.g., video or audio components) ultimately affected how much time she could spend on the design, it was evident that content was an important consideration for all three experts.

Program Settings. The experts selected for this study came from mid-sized to large state universities offering WBI courses. Anita worked with peer instructional designers, but both Catherine and Michelle were solely responsible for designing every WBI course in their programs. Still, the programs where Catherine and Michelle worked were contrastive in many ways: the e-MBA program was much smaller, cohort-based, and customized for business managers with high mobility, while the CESP handled a large volume of students and courses for cohorts of teachers pursuing their master's degrees in various education majors. Educational Outreach, where Anita worked, offered credit courses at both graduate and undergraduate levels, as well as non-credit courses. These differences affected the priorities with which each expert approached WBI design.

Steps/Procedural Knowledge in WBI Design

The steps in designing WBI courses identified by the experts remained consistent across all three cases, except for some minor details intended to accommodate case-specific requirements, such as on-site orientations or 2-week residency. When designing a new online course from scratch, all experts went through similar processes: 1) some sort of preparation to gather course information prior to meeting with faculty, 2) initial consultation meetings with faculty to learn about the course and the instructor and discuss course objectives, readings and assignments, student activities, and grading, 3) course development, where course materials were converted to Web formats, chunked and paced into weekly modules, and course logistics were determined, and 4) ongoing monitoring for course delivery and faculty training where they updated and maintained the course website and assisted with technical trouble shooting for the course throughout the semester.

Overall, Catherine and Michelle perceived the WBI design process as standardized and more reiterative than Anita did. This might be explained by the way each program was set up. Both Catherine and Michelle supported faculty at departmental or program levels, which increased the likelihood of repeatedly working with the same faculty members. Since all faculty members were from the same department or program and worked with the same cohorts of students, it was easier for both experts to recycle the WBI design process once they found a system that worked for their context.

Dissimilarly, Anita supported faculty at the university level, which means her clients might be faculty members from any discipline, and the courses assigned could be at the graduate or undergraduate level, and for credit or non-credit. Thus, she seemed to be more sensitive and accommodating to individual faculty member's needs and to have embraced more "unknowns" than the other two WBI experts.

Guidelines/Heuristic Knowledge in WBI Design

Relationships with Faculty. In all three cases, the experts underscored the importance of building a trusting relationship with faculty to succeed, and the methods or heuristics each expert used to establish rapport with them varied. Catherine took instructors to lunch as a way of building interpersonal relationship with them and presented herself more as a coach, rather than a disciplinarian, to make constructive suggestions. Michelle used frequent and direct communication as a means of building mutual trust. Anita discretely utilized her networking with other instructional designers and support staff members to learn about particular faculty members and get advice on working with "problem faculty."

The fact that all three experts identified faculty as the most influential factor when designing WBI courses and that a significant portion of heuristic knowledge elicited during the interviews was about establishing successful relationships with faculty indicates that the task of WBI design is about interpersonal relationships between the instructional designer and faculty as much as it is about instructional design and technology.

Student Workload. As Carnevale (2001) reported, students taking online courses tend to be easily overwhelmed, dealing with new technology and course interface, higher expectations/demands from the instructor, job and other commitments in addition to taking online courses, and isolation from their peers and the instructor. It was evident from interviews with all three experts that they were cognizant of this issue and developed their own heuristics to moderate student workload throughout the semester. For example, Catherine came up with faculty guidelines which suggested that students could not spend more than 20 hours a week for all three courses combined. She also advised faculty to be aware of the "flow of the semester" and consider reducing student readings and assignments before and after the intensive 2-week residency period. Similarly, Michelle shared her experience as a distance student and helped faculty establish a more realistic understanding of how much work students could reasonably handle in a single course.

Design of Course Interface. "Less is more" was a common theme across all three cases when the experts designed the interface for WBI courses. This minimalist (or functional) approach to interface design is nothing new, and Web design gurus like Nielsen (1999) and Shneiderman (1998) have long advised that the first principle of good design is to understand users' needs and try to eliminate difficulties they might encounter while navigating.

Whether the course management systems that the experts used were developed in-house or commercially purchased, the experts streamlined the course interface, deleting or inactivating all extra tools to minimize student confusion, and used only four or five basic features that were crucial to navigating the course website. For example, Catherine applied the same course design template ("course outline"-"lectures"-"discussion forum"-"assignments/grades") to every course she designed, and so did Michelle with her own selection ("content modules"-"communication tools"-"grades"-"tutorials").

Role of Instructional Designer. According to the *Chronicle of Higher Education*, many universities and colleges venturing into Web-based distance education have created instructional support units to help faculty transform their face-to-face courses into online versions, and some administrators believe that instructional designers should have some expertise in a given discipline, as well as skills and experience in project management and instructional design (2000, August 4).

Given the large volume and variety of courses they handle in practice, however, it is unrealistic to expect instructional designers to be knowledgeable in every WBI course content area they design. Among the three experts, Catherine's Hospitality and Tourism Management background was the closest match for the courses in her program (Food and Agricultural Business Management), but overall, the instructional designer's subject matter knowledge was not considered crucial to designing successful WBI courses.

The role of the instructional designer varied from one case to another, which seems to reflect how each program was set up. On one end, Michelle worked with faculty members who taught courses through the CESP, and her assistance was largely limited to creating course materials and setting up courses in WebCT. This was primarily a consequence of the sheer number of courses she handled every semester. Although she initially provided each faculty member with individual consulting, delivering a course was ultimately considered the instructor's responsibility, and once the semester began, she did not get involved with the course except for mechanical/technical trouble-shooting.

Conversely, Catherine was hands-on with faculty throughout the course development process, from advance planning to actual delivery. The e-MBA program consisted of four semester modules, and each course was an integral part of those modules. The program also had some program-specific requirements such as a 2-week residency during each semester and grading stipulations. Catherine's job was not only designing WBI courses, but also assuring that every course in the program met all program requirements and maintained consistency as a part of the whole e-MBA program package.

For Anita, course involvement fell somewhere between the other two. The CITI was loosely structured in the sense that it was not associated with a particular department or program but provided assistance to any faculty member on request. She treated each faculty member as a client and focused on his or her needs when designing courses. Technological infrastructures and staffing in each institution also seem to be a factor in determining the role the instructional designer played. The resources of the CITI team, which consisted of multimedia specialists, graphic designers, and computer programmers, as well as her peer instructional designers, enabled Anita to act as an instructional design consultant, directing her clients to various delivery options to best suit their instructional needs.

Regardless of differences, when designing WBI courses, all the experts demonstrated the ability to use heuristics to achieve the following goals: 1) to recognize and flexibly cope with contextual constraints to meet the needs of faculty, students, and the program/unit; 2) to best utilize available resources (e.g., technology infrastructures, campus-wide support structures); 3) to draw from their unique backgrounds (e.g., corporate backgrounds, experience as a distance student) as well as their training as instructional designers; 4) to maintain integrity and rigor in their courses by assuring interactivity and quality control; and 5) to plot and pace content and activities according to the "flow" of courses.

Reflections on Findings

The HTA applied to this study yielded both procedural and heuristic knowledge that guided the three experts' performance in WBI design. When we further analyzed and categorized the elicited knowledge by topic areas, however, we found relatively few heuristics related to instructional design compared to other areas such as interpersonal relationships, project management, and program/curriculum design. There may be several explanations for this lack of instructional design principles and guidelines emphasized by the participating experts, and this section approaches this issue from three different angles: 1) selection of experts; 2) implications of using course management systems in WBI design; and 3) contextual factors that constrained the WBI design process in each case.

Selection of Experts. The type and significance of knowledge that can be elicited with the HTA may vary depending on the areas and depths of expertise of participating experts. Thus, the selection of qualified experts appears to be crucial to ensure the integrity and usefulness of elicited heuristics as outcomes. The HTA does not address this issue in enough detail, however, and the biggest challenge for conducting the HTA in this study was related to selecting "expert" instructional designers, and more specifically, what selection criteria to use and how to recruit the experts.

Although there have been several research studies attempting to systematically investigate expert performance in instructional design (Rowland, 1992a and 1999; Achison, 1996; Le Maistre, 1998), few of them provided details and guidance concerning the selection of experts. Since the widely accepted "10-year rule" could not be applied due to the immaturity of the field of WBI design, we had to come up with our own set of qualifications to distinguish "expert" instructional designers from less experienced designers.

The participating experts for this study were selected based on the pre-defined criteria (i.e., relevant knowledge and experience, peer recognition, and availability), and they can be described as experienced instructional designers representing current best practice in WBI design. Still, there is no evidence to systematically verify the quality of their WBI courses in terms of pedagogical value, which may be all the more suspect due to the lack of their emphasis on instructional design guidelines during the interviews.

Case-specific Constraints. Through the three cases studied, it was clear that the work settings where the experts practiced WBI design had an impact on how the experts approached the WBI design, and more specifically, each case had context-specific constraints that hindered the experts from practicing more creative and reflective design processes. For example, the limited time and the large volume of courses assigned to Michelle each semester forced her to spend most of her time and effort on the production and management side of the WBI design process, leaving her little room for focusing on pedagogical decisions. Compared to Michelle, both Catherine and Anita had more flexibility and resources to work on individual courses. However, the tightly laid-out curriculum and the demand from the program to satisfy a particular clientele (i.e., full-time working professionals taking three courses per semester) kept Catherine from exploring diverse instructional means for creating unique and customized learning experiences in each course and forced her to follow a rather standardized process. Similarly, the types and depth of instructional activities that Anita could incorporate when she designed self-directed, non-credit WBI courses were somewhat limited because of the program's emphasis on content/information, rather than on activity/learning.

Implications for WBI Design

The use of Web-based technologies in distance education is burgeoning, and the practice of professional instructional designers assisting faculty in converting their classroom teaching into online courses is an important part of this development. Still, there are few studies that closely examine the everyday practices of instructional designers and identify their expert knowledge to benefit a broader interested audience (e.g., less experienced instructional support staff, or faculty members who want to develop WBI courses on their own) (Carnevale, 2000; Julian et al., 1999; Le Maistre, 1998). The findings of this study contribute to the body of knowledge in WBI design and untangle some of the myths and misunderstandings about the practice of instructional designers in higher education.

In earlier development of the WBI design field, researchers and practitioners raised their concerns that the new trends and practices in WBI design might inhibit, rather than promote, good education by interfering with faculty independence in developing and teaching courses, unbundling the faculty role into different (and often less qualified) personnel in developing the curriculum, teaching the course, and assessing student outcomes (Carr, 2000; Feenberg, 1999; Schifter, 2000).

However, the findings from these three cases indicate that working with a professional instructional designer improved the integrity and quality of WBI courses in at least three ways: first, faculty members went through timely, rigorous course planning to re-examine their course preparation and delivery for online teaching; second, they often benefited from having someone with more experience with distance students and the program and thus were able to build their course activities on students' previous knowledge and experience; and last, the instructional designer provided them with ongoing technical support and often worked as an arbitrator between faculty and students, enabling faculty to concentrate on their teaching rather than course maintenance and technical trouble shooting.

Implications for the HTA

Overall, the HTA interview protocols worked well in the sense that we could elicit major steps of WBI design and heuristic knowledge with which the experts performed each step. At the same time, we also faced challenges when trying to identify the simplest and most representative WBI design case.

As Chipman and his colleagues argue (Chipman, Schraagen, & Shalin, 2000), most tasks involve both observable (or "procedural") and tacit ("heuristic") knowledge. Arbitrarily separating one type of knowledge from the other is unlikely to enhance our understanding to improve task performance. Thus, during the interviews, we addressed both procedural and heuristic elements of WBI design and found that a significant portion of heuristic knowledge was triggered when associated with procedural elements.

All three experts seemed to have difficulties identifying the simplest yet representative case in WBI design. They found it difficult to cover enough depth and breadth of heuristic knowledge based on a single case. Thus, it was useful to ask them to think about ways the specific case fell short of how it should have been done (an ideal case for this version of the task) and to have them offer guidelines for how this specific case should have been done.

The HTA suggests that the task analyst interview multiple experts and identify the simplest version of the task based on consensus among the experts (Lee & Reigeluth, in press). However, it was difficult to have the three experts all think of the same case because their realms of knowledge did not overlap one another due to the differences in their program settings (e.g., student base, content areas, and resources). Thus, each expert was asked to identify the simplest version of a WBI design case that represented her work regardless of the versions that the other experts identified, and we treated each case as a separate version. It was also beneficial to periodically ask each expert questions about the chosen case, to keep the analysis focused on the flow of that version of the task.

Previous studies on the HTA (Reigeluth, Lee, Peterson, & Chavez, 2000) suggested the use of supporting materials to help experts recall details about the task performance for use as points of reference. However, all WBI courses developed by the three experts were housed within course management systems and were password-protected. To preserve the instructor's intellectual property and anonymity, we had limited access to course websites and were able to see the materials only when the experts were present to provide a password for their courses. Since we could not take anything away (e.g., capturing screen shots), asking content-specific questions or referring to a certain section of the course outside of in-person interviews was a challenge. Free access to the course websites would have been helpful to save interview time and effort as well as to increase accuracy in communicating with the experts over the phone or e-mails.

Furthermore, given the collaborative nature of WBI design, selecting faculty members who worked with the experts and conducting interviews with them might have been helpful in eliciting more detailed, authentic heuristic knowledge. Also, additional guidance for interviewing experts while they interact with faculty might be beneficial.

References

Atchison, B. J. (1996). *Roles and competencies of instructional design as identified by expert instructional designers*. Unpublished Doctoral Thesis, Wayne State University, Detroit, Michigan.

Benbenishty, R. (1992). An overview of methods to elicit and model expert clinical judgment and decision making. *Social Science Review, 66*(4), 598-616.

Carnevale, D. (2000, August 4). Turning traditional courses into distance education: Instructional designers translate professors' teaching styles into electronic content. *The Chronicle of Higher Education*.

Carnevale, D. (2001, July 17). Logging in with: Online instructor cautions against having too many activities. *The Chronicle of Higher Education*.

Carr, S. (2000, June 9). Faculty members are wary of distance-education ventures. *The Chronicle of Higher Education*.

Carr-Chellman, A., & Duchastel, P. (2000). The ideal online course. *British Journal of Educational Technology, 31*(3), 229-241.

Chipman, S. F., Schraagen, J. M., & Shalin, V. L. (2000). Introduction to Cognitive Task Analysis. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive Task Analysis* (pp. 3-24). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cuban, L. (1988). *Teachers and machines: The classroom use of technology since 1920*. New York: Teachers College Press.

Ericsson, K. A., & Charness, N. (1997). Cognitive and developmental factors in expert performance. In P. J. Feltovich, K. M. Ford, & R. R. Hoffman (Eds.), *Expertise in context: Human and machine* (pp. 3-41). Menlo Park: CA: American Association for Artificial Intelligence.

Feenberg, A. (1999). *Distance learning: Promise or threat?* Available: http://www-ohan.sdsu.edu/faculty/feenberg/TELE3.HTM [2002, February 26].

Gordon, S. E., & Gill, R. T. (1997). Cognitive Task Analysis. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 131-140). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting knowledge from experts - A methodological analysis. *Organizational Behavior and Human Decision Processes, 62*(2), 129-158.

Julian, M. F., Larsen, V. A., & Kinzie, M. B. (1999). *Compelling case experiences: Challenges for emerging instructional designers*. Paper presented at the National Convention of the Association for Educational Communications and Technology, Houston, TX.

Kitto, C. M., & Boose, J. H. (1987). *Choosing knowledge acquisition strategies for knowledge acquisition tasks*. Paper presented at the WESTEX-87 Western Conference on Expert Systems.

Le Maistre, C. (1998). What is an expert instructional designer? Evidence of expert performance during formative evaluation. *Educational Technology Research and Development, 46*(3), 21-36.

Lee, J. Y., & Reigeluth, C. M. (In press). Formative research on the Heuristic Task Analysis (HTA) process. To appear in *Educational Technology Research & Development, 51*(4).

Lohr, L. (1998). *Using ADDIE to design a Web-based training interface*. Paper presented at the SITE 98: Society for Information Technology & Teacher Education International Conference, Washington, DC.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass Inc., Publishers.

Nielsen, J. (1999). *Designing web usability*: New Riders.

Olson, J. R., & Rueter, H. H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems, 4*, 152-168.

Pearl, J. (1984). *Heuristics: Intelligent search strategies for computer problem solving*. Reading: MA, Addison-Wesley Publishing Company.

Reigeluth, C. M. (Ed.). (1999). *Instructional design theories and models* (Vol. II). Hillsdale, NJ: Lawrence Erlbaum.

Reigeluth, C. M., Lee, J. Y., Peterson, B., & Chavez, M. (2000). *Formative research on the Heuristic Task Analysis process*. Paper presented at the Annual Convention of the American Educational Research Association, New Orleans, LO.

Rowland, G. (1992a). Educating the reflective designer. *Educational Technology, 32*(12), 36-44.

Rowland, G. (1992b). What do instructional designers actually do? An initial investigation of expert practice. *Performance Improvement Quarterly, 5*(2), 65-86.

Ryder, J. M., & Redding, R. E. (1993). Integrating Cognitive Task Analysis into instructional systems development. *Educational Technology, Research and Development, 41*(2), 75-96.

Schank, R. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*: Cambridge University Press.

Schifter, C. C. (2000). Faculty motivators and inhibitors for participation in distance education. *Educational Technology, 40*(2), 43-46.

Shadbolt, N., O'Hara, K., & Crow, L. (1999). The experimental evaluation of knowledge acquisition techniques and methods: History, problems, and new directions. *International Journal of Human-Computer Studies, 51*, 729-755.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.). Reading: MA: Addison-Wesley.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394-403.

Yi, L. (1992). Knowledge acquisition - Issues, techniques and methodologies. *Data Base, 23*(1), 59-64.

Zhao, J. J., & Parks, C. (1995). Self-assessment of communication behavior: An experiential learning exercise for intercultural business success. *Business Communication Quarterly, 58*(1), 20-26.

Appendix. Summary of the Heuristic Task Analysis Method

Phase I. Prepare for Analysis and Design

1. Decide on a task to analyze and be clear about your reasons for analyzing it.
2. (Make sure you (the analyst)) have enough task knowledge to have a good command of terminology and key ideas.
3. (Make sure you) have enough knowledge about the intended uses of the task description.
4. Arrange to interview multiple experts.
5. Establish rapport with, and introduce the HTA process to, the first (and each additional) task experts.
6. Prepare for the interviews.

Phase II. Identify the Simplest Version of the Task

7. Identify the simplest version. Hold a focus group interview with multiple task experts, and help them to reach consensus about the simplest version of the task that is fairly representative of the task as a whole. Also help them to describe the conditions that distinguish that version from all other versions.
8. Analyze the heuristic knowledge. With the least experienced expert that you have not yet interviewed, analyze the heuristic knowledge (mostly principles, causal models, and guidelines) for this version of the task.
   - 8.1 With the task expert, review a record (or any other visual aid) of this version of the task.
   - 8.2 Ask the task expert to think of and describe one specific performance of this version of the task to focus on for your analysis, or ask if a videotaped performance would be a good case for you to focus on with the expert during the analysis as a form of case-based reasoning.
   - 8.3 Decide whether to use a top-down or bottom-up approach to analyzing the knowledge upon which the expert's performance is based). If top-down, use Step 8.4 and skip Step 8.5. If bottom-up, skip Step 8.4 and use Step 8.5.
   - 8.4 If top-down approach, start by identifying the general categories of knowledge that the expert uses, then proceed to analyze each.
   - 8.5 If bottom-up approach, ask the expert to describe each decision that s/he made and the process through which s/he went to make each decision.
   - 8.6 Ask the task expert to think of similar performances of the task that are within the realm of the version of the task you are currently analyzing. Use each such performance (case) to broaden the steps, guidelines, explanatory models, descriptive models, and metacognitive/decision rules so that they represent the knowledge bases the expert uses to deal with all performances for that version of the task.
9. Repeat this entire process (Steps 5 - 8.6) with the next least experienced task expert to identify any alternative views of the task and the knowledge that underlies its performance.

Note: For more detail, see Lee & Reigeluth (In press).

ASSIST-ME Model for Web-based Instruction

Hongqin Li
Franklin R. (Bud) Koontz
The University of Toledo

Abstract

This presentation will introduce an instructional design model for Web-based Instruction that may be used by faculty who teach instructional design courses and professional instructional designers and teaching faculty who design instruction for on-line courses. The ASSIST-ME Model, based on a current Delphi research study, will offer an instructional design procedure specifically intended for the uniqueness of on-line courses.

New Wine in New Wineskins

Let's take a fun quiz. After all, educators like a challenge. Choose the <u>best</u> answer from the multiple-choice question.

What would happen if you would ferment *new* wine in an *old* wineskin?

A. The new wine would take on the flavor of the old wineskin.
B. The old wineskin cannot be found.
C. The wine would not totally ferment.
D. The new wine would burst the old wineskin.

What was your answer?

If you selected A, *the new wine would take on the flavor of the old wineskin*, you are correct but it is not the best answer. The old wineskin would flavor or taint the new wine. If a different flavor of wine would be fermented, the old wineskin would prevent the new flavor from developing due to the residue remaining in the old wineskin. Besides, life is too short to drink bad wine.

If you selected B, *the old wineskin cannot be found*, you are again correct, but again, B is not the best answer. Traditionally, when wine was emptied from the original wineskin it was discarded and not used again. It had served its purpose in fermenting new wine and now the old wineskin was basically worn out.

If you selected C, *the wine would not totally ferment*, you are again correct, but it is still not the best answer. It is true that new wine could be poured into the old wineskin, but it would not totally ferment because it has already been used in the fermentation process and probably very shabby at this point.

Well now, the only remaining answer for you to select is D. *The new wine would burst the old wineskin.* During the fermentation process, certain gases are released expanding the wineskin to its absolute limit. If the old wineskin were to be used for a second time, the wineskin could no longer be stretched and would burst. And besides, who wants to lose all of that new and good wine!

Metaphors are a great way to teach and we can use a figure of speech or a phrase that designates one concept and apply it to another. The metaphor of pouring new wine into a new wineskin is illustrative of what is needed in Web-based Instruction. Traditionally we have used general classroom based instructional design models to design new Web-based instruction. This is the same as pouring new wine in old wineskins. The same, and sometimes worn out instructional strategies are still being used for a new type of instruction that we call Web-based Instruction. It is time that the new instruction for Web-based courses must be placed into a new design model that can meet the new instructional needs of our students.

Background

The designing of on-line courses requires a radical change in thinking by the instructional designer in the way the instruction is designed and presented to the student. From classroom instruction to on-line instruction is like going from an instructional television lesson that transports the student from the classroom to a distant country to be immersed in the history, culture, customs and music to a silent movie where the student may feel lonely and isolated. The instructional designer must accommodate the student in this new on-line learning environment. It is the "instructional designer who has the primary responsibility to make sure the on-line program accomplishes the learning goals – in essence, that it teaches what it is supposed to teach" (Dempsey & Van Esk, 2002).

Courses taught in Instructional Design in the area of Instructional Technology may be found in the majority of Colleges of Education. Generally, Instructional Design is a required area of study. These courses introduce a myriad of traditional instructional design models suitable for traditional classroom instruction. To date, however, no models are available that addresses or illustrates the design strategies that may be used for Web-based instruction.

Need for Web-based ID Model

Traditional classroom design models are presently being taught and used for this new form of on-line instruction. It is still contended (Dempsey & Van Esk, 2002) that Gagne's Nine Events of Instruction (1985) and Kellers (1983) ARCS Model are sufficient to use when designing on-line instruction. These models address instructional strategies and motivational strategies, respectively. However, according to a recent study (Sun, 2001), half of the 133 instructional designers who were surveyed indicated there was a need for a specific instructional design model to be created and used for Web-based courses. A second study conducted by Li (2003) using the Delphi technique, also found a need for a specific instructional design model that would address the uniqueness of this type of instruction. The majority of the respondents to the survey in Round I indicated a need for a specific instructional design model dedicated to Web-based Instruction and that the traditional models being used did not address needs of Web-based Instruction. In addition, there is a need by teaching faculty and professional instructional designers to have a model to design Web-based courses.

According to the U.S. Department of Education National Center for Educational Statistics, over 25,000 on-line courses were offered during the 1999-2000 academic year. Faculty and administrators in higher education are feeling the pressure to produce on-line learning, sometimes in less time than might be allowed for a comparable classroom course. Fifty-eight percent of two year and 62% of four-year public institutions of high learning already offer on-line education, and these numbers are expected to be 86% and 85% by 2001 (Hodgson, 1999). Currently, The National Center for Education Statistics has indicated over 9,000 colleges and universities in the United States are offering on-line courses (NCES, 2002).

A review of a variety of on-line courses, degree programs and certificate programs revealed that many of these courses are poorly designed, pedagogically unsound, and are no more than cut and past lecture notes or textbooks onto a Web site (Schweizer, 1999). The majority of teaching faculties are designing their courses on a trial-and-error basis using the same teaching and design techniques used for their conventional classroom instruction and have no researched evidence whether their web-based courses are effective.

Why do we need to know anything about new instructional design procedures? After all, we have taught traditional classroom classes for years as well as on-line courses and our students appear to be learning. Results from test scores show that our students are learning and, for the most part, receiving fairly good grades. Why can't we continue to create our Web-based Instruction (WBI) as we always have done? How difficult and time consuming will it be to use a model? And if we use an instructional model, do we have to go through every step?

These and additional questions may be asked when a new approach has been suggested to design something we think we have already successfully done. According to the present literature (Moore & Kearsley, 1996) addressing the issue of student learning, the majority of designers frequently do not attempt evaluation of instruction during the time the instruction is being designed and very little or no evaluation of the effects of the instruction after the course has been completed. This lack of evaluation may not be completed for several reasons. It takes time to establish evaluation criteria and the evaluation instrument. It takes time to administer the instrument to the students and it takes time to interpret the results. If what we find is negative and a redesign of instruction must be conducted, additional time must be committed to design process. Time is very precious to all designers and teaching faculty. That being the case, how do we really know if our instruction is effective? Are the students learning what they are supposed to learn? And, how efficient is our designed instruction? Do the students have difficulty understanding our instruction and course assignments? Can the students maximize their learning in a minimum amount of time?

Almost any model may be used to systematically design classroom instruction. It appears, however, that traditional classroom instructional models (Li, 2003; Sun, 2001) may not meet the needs or answer many of the problems that arise in the design of on-line courses. Li (2003) found in her Delphi study that nearly 70 percent of the professional panel members agreed that there is a need for a specific instructional model to be designed and used for Web-based courses. Traditional classroom ID models are not designed for the uniqueness of distance learning, i.e., the teacher is geographically separated from the student and that the student, for the most part, interacts with the instruction independently with delayed feedback. In addition, it is also readily accepted by experienced designers that an instructional design model is more of a guide for the design process. Again, Li found nearly 68 percent of the panel members did not use all of the steps of the selected model. The designer may use his or her own judgment and omit steps to implement instruction that is needed for meaningful learning. The main objective, however, for systematically designing any Web-based instruction (WBI) should be the "first-time design" and should satisfy students' learning needs. If the "first-time design" of instruction is planned properly, design critiques will hopefully result in a positive evaluation and the designer promotes meaningful learning.

Research Design

This descriptive research study, using the Delphi technique investigated which existing traditional instructional design models were used and why some were inappropriate when used to design Web-based instruction (WBI), and which essential steps must be included in the design of Web-based Instruction. The population of this study consisted of practitioners, researchers, professors, and graduate students who were actively involved in the instructional design activities of Web-based instruction. Since Web-based instruction is a new area of study, there is no comprehensive list of such a population. This study included members of the Design and Develop Division and the Distance Learning Division of the Association for Educational Communications and Technology (AECT).

Forty-three respondents volunteered to be the Delphi panel members. Based upon the responses from the 78 early active participants, over three hundred Items were obtained as the essential steps for designing Web-based Instruction (WBI). Duplicates were dropped. The "near duplicate" items were retained as unique since they had presented a different perspective. Therefore, 101 items were generated for the round one voting. After three rounds of Delphi study, fifty-six items were retained as the most essential steps for Web-based Instruction.

Findings and Analysis

Fourteen (40%) out of the 43 people who followed models used Dick and Carey Model when designing their Web-based Instruction (WBI). Thirteen (37.14%) used ADDIE Model. The same number of people used Rapid Prototyping Model. Nine people (25.71%) used the Gagne and Briggs Model. The Smith and Ragan Model and the ASSURE Model were also among the top six (see Table 1).

Table 1 *List of Models Followed by Panel Members*

List of Models	Frequency	Percent
Dick and Carey	14	40.00
ADDIE	13	37.14
Rapid Prototyping	13	37.14
Gagne and Briggs	9	25.71
Smith and Ragan	5	14.29
ASSURE	4	11.43

Ten (23.26%) out of the 43 panel members indicated that they followed every step of the model they selected when designing Web-based Instruction. Twenty-nine Delphi panel members (67.44%) said they did not follow every step of the selected model(s). Three (6.98%) answered both "Yes" and "No" depending on the specific situations. One member did not answer this question (see Table 2).

Table 2 *Follow Every Step of the Model of Choice*

Category	Frequency	Percent
Yes	10	23.26
No	29	67.44
Yes and No	3	6.98
Not Applicable	1	2.33
Total	**43**	**100.00**

On the question whether there was a need for a specific instructional design model for Web-based Instruction (WBI), thirty 30 (69.77) answered “Yes” while thirteen (30.23%) of them answered “No” to this question (see Table 3).

Table 3 *Need for a specific instructional design model for WBI*

Category	Frequency	Percent
Yes	30	69.77
No	13	30.23
Total	**43**	**34.13**

Essential Steps for Designing Web-based Instruction

The respondents generated over three hundred items as essential steps for Web-based Instruction. Duplicate items were dropped. After that, one hundred and one items were sent out to the panel members for evaluation. Round two had 84 items including six new ones. Round three had 61 items. After three rounds, 56 items were retained, each with a mean score of 3.60 (on a 5-point scale) or higher (see Table 4).

Table 4 *Items Retained After Three Rounds (in alphabetical order)*

Rank	Item
1	Analyze Content
2	Analyze Context of Learning
3	Analyze Instructional Goals
4	Analyze Learner Interaction
5	Analyze Learner: Competencies
6	Analyze Learner: Knowledge Level
7	Analyze Learning Tasks
8	Analyze Needs
9	Analyze Objectives
10	Assemble All Elements
11	Assessing Technology Available to Students
12	Conduct Frequent In-process Review of Project
13	Conduct Quality Assurance
14	Conduct Summative Evaluation of Student Learning
15	Create Assessment Instruments
16	Create Interaction
17	Create Media Elements

18	Create Script and Storyboard
19	Define Course Outline/Module Breakdown
20	Define Test Plans: Formative and Summative
21	Design Activities that Will Lead to Outcomes
22	Design Course Goals
23	Design Instruction Including Strategy
24	Design Interactions
25	Design Interface
26	Design Learning Activities
27	Design Learning Outcome
28	Detailed Design of Content
29	Determine Cross-platform Issues and Limitations
30	Determine Delivery Strategies
31	Determine Media Requirements
32	Develop Content Materials
33	Develop Instructional Strategy
34	Develop Specific Outcome Objectives
35	Development (SMEs)
36	Formative Evaluation
37	Front-End Analysis with Faculty Members
38	Goal and Evaluation Alignment
39	Have End-Users/Learners Evaluate the Instruction and Revise
40	Identify Interactive Elements
41	Identify Learning Activities
42	Identify Learning Strategies
43	Identify Media Needs
44	Implementation of Course
45	Implementation of Technical Support
46	Message Design
47	Navigational Structure Identified
48	Perform Graphic User Interface Design
49	Practice Items/Exercise/Online Quiz Development
50	Provide Technical Support
51	Review Interaction (student, instructor, group) Design
52	Review Site with Instructor
53	Test and Revise Instruction
54	Test and Revision
55	Test Prototype
56	Use Feedback of Course Units to Develop Whole Course

The 56 items that were retained after the three rounds of evaluation by the Delphi panel were essential steps in designing Web-based Instruction. From this list identified by the Delphi panel, it appeared that there were specific items that could be categorized and placed into a new ID model specifically dedicated to Web-based Instruction (WBI).

To design this model, a careful analysis was made of the 56 items. These items were then grouped into major categories that contained similar functions. Major categories were developed, specifically: Analysis, State Performance Objectives, Select Instructional Materials, Organization Content, Teaching Strategies, and Media, Implementation, Solicit Student Response, Test, Evaluate, and Revise Instruction, and Maintenance. The major categories became an acronym for the model called ASSIST-ME model (see

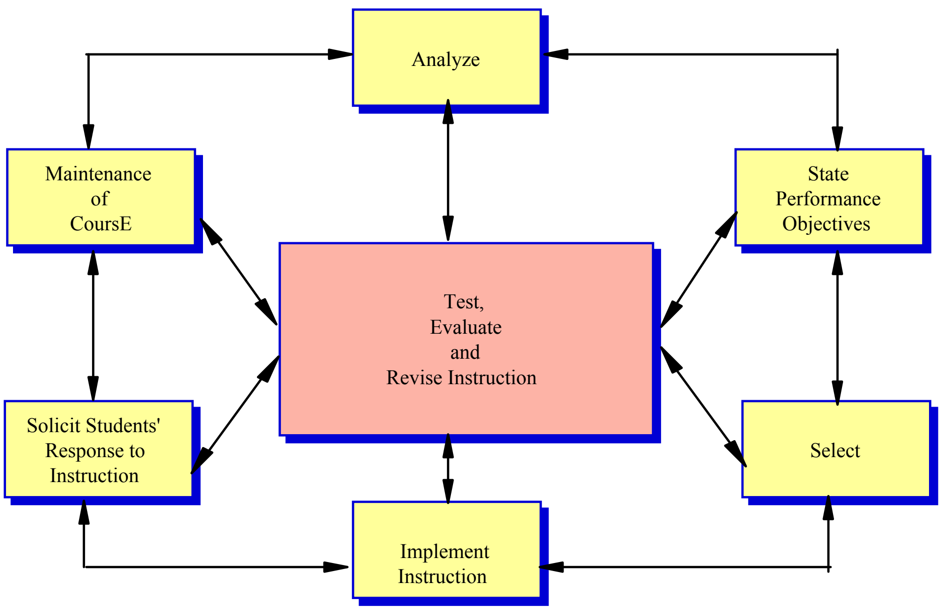
Figure 1). With each of these categories, subparts were identified.

When the elements were assigned to specific design categories, some items were deleted due to complexity, the length of time it would take to complete and the limited value it would add to the design process. Judicious decisions were made as to which of the 56 items should be included.

There appeared to be a very heavy emphasis on testing, evaluation, and revision of instruction both in formative as well as summative evaluation. Therefore, the model was designed with an emphasis on this procedure.

The following model has been designed by establishing the major categories and selecting subparts from the 56 items (see Figure 1).

Figure 1. ASSIST-ME Model for Web-based Instruction



The linear ASSIST-ME Model with sub-steps:

Analyze
- Instruction
  - Goals
  - Domains of learning
  - Knowledge Objectives
- Instructional Setting
  - Prerequisites
  - Materials Needed
  - Relationship to other Courses
- Students
  - Student Demographics
  - Specific Entry Skills
  - Learning Styles

State Performance Objectives
    Class and Level
    Performance Skills Required
    Conditions/Givens
    Degree/Criteria

Select
    Instructional Materials
    Organize Content of Lesson
    Teaching Strategies
    Media

Implement Instruction
    Course
    Individual Lessons

Solicit Students' *Response* to Instruction
    Case Studies
    Investigations
    Debates
    Interactive Discussions
    Simulations

Test, Evaluate and Revise Instruction
    Lesson Content
    Teaching Strategy
    Media
    Student feedback

Maintenance of CoursE
    Design Interface
    Identify Navigational Structure
    Graphic Design
    Technical Support

The purpose of this presentation is to introduce an instructional design model for Web-based instruction that may be used by faculty who teach instructional design courses and professional Instructional Designers and teaching faculty who design instruction for on-line courses. The ASSIST-Me Model, based on a current Delphi research study, *Investigation of an Instructional Design Model for Web-based Instruction (WBI)* (Li, 2002), will offer an instructional design procedure specifically intended for the uniqueness of on-line courses. Data obtained from a panel of professional instructional designers were synthesized into an instructional design model that contains essential elements to be included in the design process. The presentation will demonstrate a step-by-step procedure that suggests how the instruction may be designed.

The ASSIST-Me Model for WBI will be presented as an open system, i.e., once the Analysis Phase has been completed (see ASSIST-Me Model), the designer may begin to design other parts of the instruction and will not be forced to follow a linear system. The model will give the designer maximum flexibility when creating instruction.

References

Dempsey, J. V., & Van Eck, R. N. (2002). Instructional Design On-Line: Evolving Expectations. In R. A. Reiser & J. V. Dempsey (Eds.). *Trends and Issues in Instructional Design and Technology* (pp. 281 – 294). Pearson Education, Inc., Upper Saddle River, New Jersey.

Gagne, R. M. (1985). *The conditions of learning and theory of instruction* (4th ed.). Fort Wroth, TX: Holt, Rinehart and Winston.

Hodgson, P. (1999). How to teach in cyberspace. *Techniques*. *74*(5), 34 – 36.

IPEDS College Opportunities On-Line. (ND). National Center for Educational Statistics. Retrieved January 28, 2003 from http://nces.ed.gov/ipeds/cool/

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.). *Instructional design theories and models: An overview of their current status* (pp. 386-434). Hillsdale, NJ: Lawrence Erlbaum Associates.

Li, H. (2003) *Investigation of a New Instructional Design Model for Web-based Instruction: A Delphi Study*. Unpublished doctoral dissertation. The University of Toledo. Toledo, Ohio.

Moore, M. G. & Kearsley, G. (1996). *Distance Education: A Systems View*. Belmont, Ca: Wadsworth Publishing Company.

Schweizer, H. (1999). *Designing and Teaching an On-line course*. Needhaven Hights, MA. Allyn and Bacon.

Sun, X. (2001) *An Investigation of Instructional Design Models for Web-Based Instruction*. Unpublished doctoral dissertation. The University of Toledo. Toledo, Ohio.

Exploring Small Group Interactions in Computer Mediated Communication

Yi-Mei Lin
Guan-Yu Lin
Hsinyi Peng
James Laffey
University of Missouri-Columbia

Abstract

This study investigated a method for visually mapping student actions and interactions during collaborative learning. The method has potential for articulating patterns of activity and may help us better understand how teams of students use computer mediated communication (CMC) for learning. While the primary interest in this study was to advance methodology for articulating interaction patterns in CMC, the presentation will describe the interaction patterns that emerged in group discussion boards for assignments in an online graduate course.

Introduction

This study investigated an approach to advance our understanding of small group interaction by articulating the patterns of interactions during two learning activities in an online graduate course. From the sociocultural viewpoint, cognition developed when interacting with people and the social context. Vygotsky (1978) believes that through engaging activities social interaction and collaboration with others help individuals make progress through their zone of proximal development (ZPD) (Oliver, Omari, & Herrington, 1998). Consequently, research supports that collaborative activities help students develop higher-level reasoning (critical thinking) and meta-cognitive thought (Johnson & Johnson, 1989; Johnson & Johnson, 2002).

During the last decade, computer-mediated communication (CMC) has been shown as an innovative tool to support learning in education. Online discussion boards/forums tend to be the most-frequently used feature in CMC; thus, it is important to investigate online asynchronous interaction so as to identify approaches that facilitate collaboration and learning. Previous research in CMC focused on the frequency and word counts of contributed messages (Mason, 1991) provided too little information about the complexity of interaction and learning occurred in online discussion. Counting the quantities of discussion must be accompanied by analyzing the content and patterns of discussion. As a result, there is a growing research interest in applying content analysis, a research technique involving both numeric and interpretive data analysis, to analyze the contents and interactions of online discussion boards (e.g., Hara, Bonk, & Angeli, 1998; Henri, 1992; Hillman, 1999; Howell-Richardson & Mellar, 1996; Kanuka & Anderson, 1998; Mowrer, 1996). With anticipation, a more comprehensive picture of online discussion can be understood through using both quantitative and qualitative research measures (Hara, Bonk, & Angeli, 1998). The primary purpose of this study was to explore how students interact on discussion boards during small group work and to articulate the patterns of interaction.

Methodology

Context and Participants

This study examined the third and fourth weekly activities in an online graduate level course delivered through Shadow netWorkspace™ (SNS) during the Fall of 2002. SNS is a network-based learning system where users have a set of tools for representing, organizing, sharing, communicating, and collaborating in processes related to learning. During the activities eighteen students were randomly assigned to small groups of four or five members. The students were encouraged but not restricted to use the tools in SNS to work as a group. Each weekly activity required a series of learning sub-acts including reviewing instructional materials and producing a group report to answer four questions.

Instrument

Because the structure of interaction patterns among students and the nature of the interaction itself were very complex, narrative descriptions of these interactions were very difficult. Visual mapping (interaction maps) of inter-message referencing for each week of each group was done to represent the actions and interactions during each group's weekly learning activity. After interaction maps were produced, messages were coded into two dimensions: social aspect of task and content aspect of task. The inter-rater reliability was .96, as quantified by Holsti's (1969) coefficient of reliability (C.R.). C.R. is the most commonly used statistic method of reporting inter-rater reliability in which the number of agreements between rater 1 and rater 2 are divided by the total number of coding decisions made by both raters. Social aspect of task messages include announcement (e.g., "I already started a group…"), collaborative strategy (e.g., "Why don't we try shadow talk for synchronous conversion?"), greeting (e.g., "Hello team members!"), compliments to others (e.g., "Well you all have done a nice job of answering this one!") and completion of the group work (e.g., my two cents). In contrast, content aspect of task messages focus on task discussions. The following paragraphs were examples of content aspect of task messages:

> *"I think that the articles of association are best described as principles because they are the important and driving ideas of the institute. They are almost like a mission statement or a vision statement."*

> *"Standard practices help the designer because they provide methods for creating interfaces that users are most comfortable and familiar with. If designers create similar interfaces, users will be able to easily transfer their skills from one product to another because they know what to expect."*

The messages of each learning activity were plotted chronologically and hierarchically to represent the actions and interactions within each group. Thus the interaction maps were plots of the types and distribution of messages for each week and each group. Similar to the technique of concept mapping, interaction maps show the connections and relationships between messages (Figure 1). Messages are represented as oval or cloud nodes in an interaction map. An oval node is used to represent a content aspect of task message, while a cloud node indicates a social aspect of task message. A rectangle represents the final group report submitted for grading. A link between messages indicates how one message refers to anther. For example, A -> B means Student A responds to Student B directly. A linear-dotted line indicates a summarizing within a single thread, and an S-dotted line represents an indirectly summarizing action across threads. Finally, an isolated oval or cloud node indicates an attempted thread with no further discussion. After analyzing these visual maps, several unique patterns of the interaction among small groups emerged.

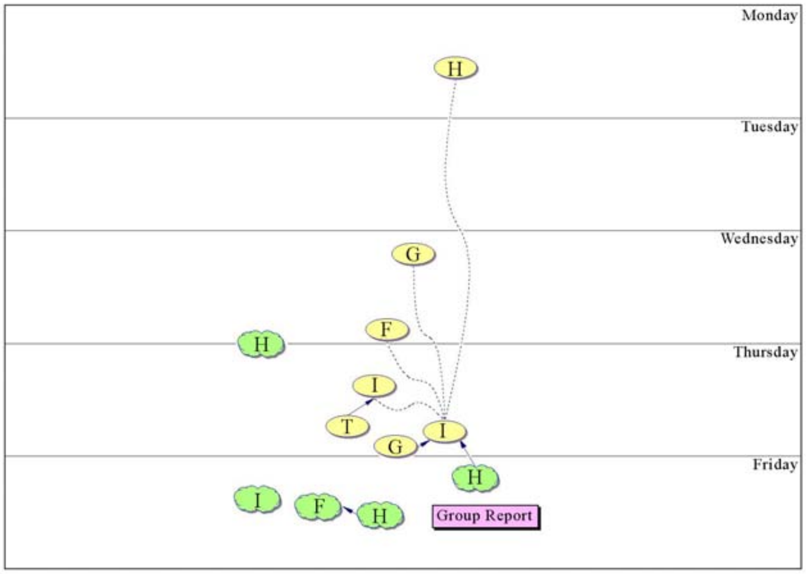


Figure 1. Interaction Map of Group 2 in the Fourth Weekly Activity

Preliminary Findings

The patterns of interactions shown by each team are different. However, these patterns share a common trend in that more interactions took place when team members exchanged social aspect of task messages. We found that these social aspect of task messages were primarily used to discuss collaborative strategies such as negotiations for utilizing appropriate tools to best synthesize final reports. Once team members reached their agreement, they started focusing on answering the questions the instructor gave to them. Content aspect of task messages were dominate unless new strategies were proposed or the group encountered technical problems. This result confirms other studies that report more task-oriented communication in CMC when compared to face-to-face communication (Walther, 1992; Jonassen & Kwon, 2001). For instance, in Group 2, the amounts of content aspect of task messages in both weekly activities were approximately equal. The team members answered these questions individually, and one member volunteered to summarize the individual answers and write a group proposal. In contrast to the nearly equal number of content aspect of messages in both activities, only five social aspect of task messages were generated in the fourth activity, while there were twenty-three social aspect of task messages in the third activity. This result may be due to this group adopting a collaborative strategy in the third week that carried over to the fourth week. In contrast Group 3 utilized different strategies between the third and fourth activities. Consequently, social aspect of task messages occurred again in the first few days before the members started posting their individual work (See Figure 2 and 3). More findings about the patterns of interactions, the use of SNS tools, and the implications of using interaction maps will be discussed in the presentation.

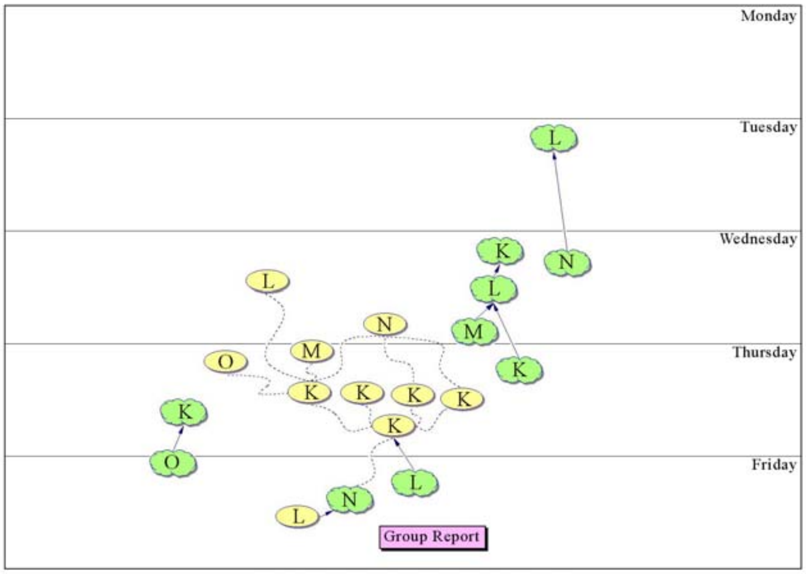


Figure 2. Interaction Map of Group 3 in the Third Weekly Activity

Figure 3. Interaction Map of Group 3 in the Fourth Weekly Activity

Conclusion

This study advanced a method for visually mapping student actions and interactions during collaborative learning assignments. The method can be used to show how students use CMC for learning activities and for deriving improvements in the CMC systems and pedagogical strategies.

References

Blake, C. T. & Rapanotti, L. (2001). Mapping interactions in a computer conferencing environment. *Proceedings of EuroCSCL2001, Maastrict.*

Hara, N., Bonk, C. J., & Angeli, C. (1998). *Content analysis of online discussion in an applied educational psychology* (Tech. Rep. No. 2-98). Bloomington, IN: Indiana University, Center for Research on Learning and Technology.

Henri, F. (1992). Computer conferencing and content analysis. In A. R. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp. 115-136). New York, NY: Springer.

Hillman, D. (1999). A new method for analyzing patterns of interaction. *The American Journal of Distance Education, 13*(2), 37-47.

Holsti, O. (1969). *Content analysis for the social sciences and humanities.* Don Mills: Addison-Wesley Publishing Company.

Howell-Richardson, C. & Mellar, H. (1996). A methodology for the analysis of patterns of participation within computer mediated communication courses. *Instructional Science, 24,* 47-69.

Johnson, D. W., & Johnson, R. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (2002). Cooperation and the use of technology. In D. Jonassen (Ed.), *Handbook of Research on Educational Communications and Technology*. In press.

Jonassen, D. H. & Kwon, H. (2001). Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research & Development, 49*(1), 35-51.

Levin, J. A., kim, H., & Riel, M. M. (1990). Analyzing instructional interactions on electronic message networks. In L. M. Harasim (Ed.), *Online education: Perspectives on a new environment* (pp.185-214). New York, NY: Praeger.

McDonald, J & Gibson, C.C. (1998). Interpersonal dynamics and group development in computer conferencing. *The American Journal of Distance Education, 12*(1), 7-25.

Mason, R. (1991). Methodologies for evaluating applications of computer conferencing. In A. R. Kaye (Ed.), *Collaborative Learning through Computer Conferencing.* Heidelberg: Springer-Verlag.

Mason, R. (1998). The globalisation of education: Synchronous versus asynchronous distance education. Retrieved March 10, 2003, from http://iet.open.ac.uk/pp/r.d.mason/globalbook/syncasync.html

Mowrer, D. E. (1996). A content analysis of student/instructor communication via computer conferencing. *Higher Education, 32,* 217-241.

Oliver, R., Omari, A., & Herrington, J. (1998). Investigating implementation strategies for WWW-based learning environments. *International Journal of Instructional Media, 25(2),* 121-138.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Walther, J. B. (1992). Interpersonal effects in computer-mediated communication: A relational perspective. *Communication Research, 19*(1), 52-90.

Sense and Nonsense in Online Interactions: Can Collaborative, Non-Informative Exchanges Build a Community of Practice?

Cindy Hui Lio
Joan Mazur
University of Kentucky

Abstract

This paper describes preliminary findings regarding how seemingly nonsensical conversations in a synchronous Book Talk among reading teachers might, in fact, support the development of a community of practice. Using Burnett's (2000) typology of exchanges for virtual communities, noninteractive and interactive online behaviors, as evidenced in the text-based 'talk', are explored. An analysis of conversation showed evidence of mutual engagement, joint enterprise and the development of shared repertoires, three activities related to the maintenance of coherence in communities of practice.

Introduction

Today's prevalent practice of online chats and discussion forums through the World Wide Web dramatically transforms the technology landscape in relation to human behaviors. The widely used Internet gives rise to a social interaction phenomenon in which participants 'converse' using electronic textual messages as their means of communication. These emergent computer-mediated environments are more than merely as a source of information, entertainment, and basic communication. In some respects, the Internet becomes an alternative avenue for people's self-expression, stress release, professional development, search of identity, socialization, and social learning (Burnett & Bonnici, 2003; Turkle, 1995).

By definition, online communities are "designed communities using current networked technology, whereas communities of practice emerge within the designed community via the ways their participants use the designed community" (Johnson, 2001, p. 45). Within the context of this definition, the term online communities covers the broad spectrum of both the synchronous and asynchronous forums such as web-based newsgroups, discussion boards, listservs, conferences, chat rooms and so forth. Regardless of the format or nature of these online communities, the pervasive underlying commonality among computer-mediated communities is the social dimension of communication. The complexity of communication in online communities characterized by the void of face-to-face interactions and text-only discourse provides a fertile ground for the study of human interactions in these evolving virtual communities of practice. Much recent study on virtual communities has been devoted to topics on gender-related issues, online participants' communication styles, potentials and challenges of online communities within the realm of education (Ferry, Kiggins, Hoban, & Lockyer, 2000; Kreijns, Kirschner, & Jochems, 2002; Lapadat, 2002; Rogers, 2000; Soukup, 1999). Moreover, the impacts of online communities on the society as a whole, and the social interaction and communication of related issues are reported in the literature (Agres, Igbaria, & Edberg, 1998; Cummings, Butler, & Kraut, 2002).

In studies of both face-to-face and online conversation, the social and communicative significance of seemingly mundane conversation has been well documented (Preece, 2000). However, in online discussions in which teachers engage in professional conversation, "off-topic" or "whimsical" dialogue is often viewed as being off task or wasting time. This paper examines the social aspect of a particular synchronous Book Talk among reading teachers and describes preliminary findings regarding positive effects of seemingly unrelated conversations in this environment might, in fact, support the development of particular aspects of a community of practice as defined by Wenger (1999).

Conceptual Framework of the Study

Much literature in recent years has addressed the various social aspects of a community of practice, particularly in association with the burgeoning virtual communities. Moreover, studies that analyze the actual dialogues or discourse patterns of virtual communities are diverse in scope and in theoretical framework. For example, Muramatsu and Ackerman (1998) studied the social management and activities of game participants in a combat MUD community. From a six-month engagement and observation of the selected virtual game community and from interviews with other MUD participants and administrators, the researchers found that meaning-making resulted largely from collaboration and conflicts. Their analysis of dialogues and interactions among game participants also identified the significant role of game administrators in structuring and maintaining the social activities within the virtual entertainment environment. In a longitudinal study by Schoberth, Preece, and Heinzl (2002), actual email postings were analyzed for an understanding of the social behaviors within an online community. Their quantitative analysis of three years' email messages from a German financial service provider revealed a high correlation between information load, communication strategies and communication activities. In her study of the structural and behavioral patterns of computer mediated textual communication, Herring (1999) sought explanations for the complex and frequently problematic nature of synchronous and asynchronous interactive exchanges. By analyzing samples of online dialogues, Herring (1999) found that the prevalent problems of topic decay, separated response sequence, exchange gaps and exchange overlapping resulting from the technology often created confusion and incoherence in online textual dialogues. However, her study also indicated that online participants frequently sought compensatory strategies to adapt. Furthermore, many online users extend their creativity in the textual form of expressions by playing with words and manipulating textual expressions. Although the aforementioned literature concerns social interaction and communication, these studies do not explore how the complexity and multiple dimensions of social exchanges in online conversation relate to the development of a community of practice. Toward this end, the authors adopted Wenger's (1999) definitions of community of practice as the conceptual framework to examine this potential relationship using the dialogues of a synchronous Book Talk among reading teachers.

According to Wenger (1999), a community of practice is one in which members engage in the same types of work and thereby experience a continual process of unfolding, integrating, negotiating, and development of related knowledge, expertise and support. The interwoven relationships of members in these communities of practice encompass three essential dimensions: mutual engagement, joint enterprise, and shared repertoire. Each of these dimensions has unique characteristics of interaction that maintain the coherence of the communities. For instance, mutual engagement entails the process of evolving forms of relationship and community maintenance; joint enterprise encompasses the process of understanding, accountability, and synergistic coordination; and shared repertoire embeds the process of developing styles, artifacts and discourses (see Figure 1).

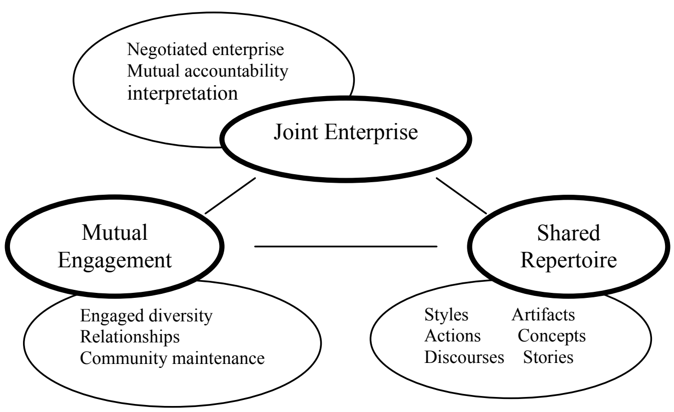


Figure 1. Dimensions of practice as the property of a community (Wenger, 1999, p.73).

As much online conversation is text-based, not surprisingly, it differs in significant ways from face-to-face communication. Burnett (2000) has developed a typology of social exchanges in virtual communities. According to Burnett (2000), online social interactions can be divided into interactive or non-interactive behaviors. Within the interactive category, online talk and behavior can be either hostile or collaboratively interactive. In turn, the collaborative interaction category is further subdivided into behaviors not specifically oriented toward information and those that are. The initial focus of this preliminary investigation of the conversation in a synchronized teacher professional development forum (chat) was on the information-oriented behavior in the interactive category. However, the authors began to question their tendencies to dismiss the joking or otherwise seemingly non-informative interactions as irrelevant. This closer look at the non-information-oriented behaviors is the focus of the study.

Methods

The Forum Description: Teachers.Net Book Talk

The archived transcript of a synchronous Book Talk session in Teachers.Net services, which is a division of Teachers Web, was the primary data source for this initial study. The motto of this teacher-oriented website is "improving education through technology and collaborative support...together we share an incredible potential to change the face of education" (Bott, 2002, 1). The site is supported by educational organizations and "commercial enterprises seeking technical and promotional solutions to researchers of cutting edge theory and applications" (Bott, 2002, 2). AIG VALIC, an organization that offers financial services and products, is the major sponsor; others include Azusa Pacific University, Apples 'N' Acorns, LASIK for Educators, Meet the Masters NRTA, ReadingLady.com, and VIF. Aside from educational advertising by the sponsors, there are numerous instruction-related chat rooms, chat boards, live meetings, and topical meeting archives. Among the diverse chat forums in Teachers.Net, the archived chat meeting appeared to be the most appropriate for the observation of online community behavior. These archives are transcripts of live events in areas such as phonemic awareness discussions or grade level teachers' chat. Generally, live chat or discussion sessions take place at 9:00 p.m. Eastern

Standard Time in the virtual Meeting Room.

In order to entice advertisers, Teachers.Net reported on their general demographics obtained during a seven-day data collection from 1200 voluntary participants taken in the fall of 1998. According to the note to potential advertisers, this Teachers.Net survey claimed 91.87% participants surveyed were teachers; 90% were female. Furthermore, among the 95% of participants who reported being college-educated, 44.4 % possessed advanced degrees. The average age of the surveyed participants was 40.6 years (see Table 1).

Categories	Unit Breakdown	Surveyed Results
Age of Participants	Average age	40.6 years
Curriculum Taught	Early Childhood	16.5 %
	Elementary	54.2 %
	Middle/HS	21.9 %
	Other	7.4 %
Education	HS	0.25 %
	Some College	2.68 %
	Assoc.	1.76 %
	Bachelor	51.26 %
	Masters	41.36 %
	Doctorate	2.68 %
Ethnicity	Caucasian	93.32 %
	AA	2.23 %
	Hispanic	1.71 %
	Multi	1.63 %
	Other	1.11 %

Table 1. *Teachers.Net survey results - fall 1998*

Categories	Unit Breakdown	Surveyed Results
Sex	Female	90.7 %
	Male	9.3 %
Marital Status	Married	73.8 %
	Single	26.2 %
Occupation	Teacher	91.87 %
	Ed. Student	3.6 %
	Admin.	1.97 %
	Other	2.56 %
School Type	Public	85.2 %
	Private	12.0 %
	Other	2.8 %
Student Age	Average	8.8 years
	5-9	52.0 %
	10-14	26.1 %
	15-18	7.3 %
	Other	14.6 %
Teaching Experience	Average	13.66 years

Table 1. *Teachers.Net survey results - fall 1998*

Like all the other free Teachers.Net chats and meetings, the chosen archived session, Book Talk, met as a scheduled real-time event. Starting its first session on April 17, 1998 with seven teacher participants, Book Talk has been meeting irregularly until March 2002 and thereafter continued weekly until the beginning of 2003.

Participants

Based on the content of the Book Talk transcript, we inferred that participants were from various parts of the United States such as California, Michigan, Tennessee, Oregon, Illinois, and Virginia, and Canada. Explicit sex, race, and ethnicity information was not available. However, sex was inferred by participant names, although we had no way to confirm our inferences. Participants also did not specify grade level they taught, although again, it appeared from the topics that several were elementary teachers. Participants were given pseudonyms in the excerpts of chat logs that appear in this report.

Procedures

Access

The interest of the authors to explore the engagement of teachers in professional online discussions led to the search on teachers' discussion boards, chartrooms, and synchronous meetings. However, without official membership, not many of the teacher related websites provided easy access to the information of online discussion. The selection of Teachers.Net was largely due the ready access provided to the events and archived transcripts. For the purpose of this study, the authors chose Book Talk among other discussion sessions in the archive for the simple reason that the dialogues appeared to be substantive, lively and amusing. The authors contacted the online administrator of Teachers.Net to seek permission to use the transcriptions of Book Talk chat archives. However, we did not receive a reply. Consultation with our Institutional Review Board (IRB) resulted in exemption status for the current study because the logs of Book Talk chats are a publicly available resource and moreover had been posted on Teachers.Net before this study began.

Sample

Sampling in conversation analysis (CA) precedes from a different starting point and with different epistemological assumptions, than that of sampling in quantitative frameworks (Mazur, 2004). Drawing on the insights from naturalistic observation and biology, conversation analysis justifies the use of 'specimens' of conversation. Specimens (in biological study, for example) represent the phenomena or organism under study that are directly observable in the particular individual. While such specimen sampling may have to follow certain criteria, there is no need to sample from a random population. For example, a biologist examining a species of tadpoles would simply choose individuals from that species. Perhaps, by observing the species in some natural setting, the biologist would not choose one that seemed highly unusual, but there would be no need for a statistical sample of tadpoles since it is the category of specie, the specimen, that is of interest. Similarly, a CA study might focus on any category of talk-in-interaction such as repairs or greetings and then simply select any specimen of conversation generated in a naturalistic setting. As Hutchby (2001, p. 51) has claimed, "The logic of CA...in terms of data selection suggests that *any* [his emphasis] specimen is a 'good' one, that is, worthy of intense and detailed examination (as quoted in Mazur, 2004, p.1083)"

The Book Talk session took place on November 1, 2002 with 16 participants and 1025 exchanges. However, the observed segment has 50 exchanges among eight regular participants. The major themes of the selected discussion session center on an upcoming conference, personal features of participants such as hair color and age, school substitute issues, Halloween party, number of school days, phonics, readings, and issues related to teaching and class schedules.

For analysis purposes, the archived chat logs on November 1, 2002 were copied and formatted, which yielded 46 pages of double spaced transcripts. The dialogue of the complete discussion session, which has 1025 lines of exchanges, was numbered for the ease of analysis. Names of the participants were then italicized and were referred by pseudonyms. Finally, from the formatted transcript of dialogue, the researchers chose a specimen comprised of lines 71-121 and named it as the "Hair Talk" which enables generalization of observations on this talk-in-interaction of participants. Expecting a discussion of substance on children's book and reading in general, the researchers were initially disappointed but then intrigued by the somewhat silly repartee regarding hair color and style that ensued. What meaning could such off task online exchanges have in this self-initiated interactive context?

Gunawardena, Lowe and Anderson (1997) have cautioned against the sole use of conversation analyses to characterize the interactions in online forums. Their concern is that the focus on the details of talk-in-interaction may result in the loss of the so-called bigger picture, the meanings, of the conversation. They justify their position as follows:

> The totality of interconnected and mutually-responsive messages which make up the conference, and perhaps more: 'interaction' is the entire gestalt formed by the on-line communication among the participants. The process observed in the debate is akin to thinking on 'distributed cognitions' (Salomon, 1993, p. 256) where he states individual and distributed cognitions interact over time, affecting each other and developing from each other. (p. 407)

Heeding their concern, we included a content analysis to complement the conversation analysis of the selected specimen.

Analysis

Using the frameworks from both traditional conversation analysis (ten Have, 2000) and those that focus on virtual conversation (Herring, 1999), the Book Talk online discussion was analyzed. Turn-taking, sequencing and content of conversation were examined for normative patterns in the conversation. These patterns were then related to Burnett's (2000) typology for exchanges in a virtual community and to the concepts from Wenger's (1999) notion of a community of practice, specifically, mutual engagement, joint enterprise and the development of a shared repertoire.

Analyzing the turn-taking organization in a conversation is "one of the core ideas of the CA enterprise" (ten Have, 2000, p. 111). The basic idea is that turn-taking is the procedure that carries the conversation from one speaker to the next. During a normative conversation, the speaker selection, the tone of speaking, and the gaps in between responses are just some of the phenomena of turn-taking organization. Since online written dialogues obviously do not exhibit the tone of speaking as in face-to-face communication, expressions such as exclamation points and smiling symbols (emoticons) are often used instead. With speaker selection, the three prevalent ways that speakers exhibit their turn-taking during a conversation are other-selection, self-selection, and continuation (ten Have, 2000). Each turn-taking option illustrates the participants' actions within the context of the conversation.

Findings

Turn-Taking

Excerpt 1

71.	*Terri*	(((Annie)))
72.	*Todd*	Hi Annie!!!!!!!!
73.	*rita2a*	(((Annie)))
74.	*amy*	Hi Annie!
75.	*kevin*	Hey Annie!!!
76.	*AnnieL*	Hi Everyone!
77.	*amy*	Todd, why will you have small animals in your hair?

Continual and self-directed turn-taking was evident throughout the selected conversation specimen as demonstrated by the greetings of Terri, Todd, Rita, Amy, and Kevin (see Excerpt 1, 71-75) to Annie. No one in particular suggests them to welcome Annie. In response, Annie does not hesitate in returning the greetings. Line 77 (see Excerpt 1) typifies an other-directed organization and self-selecting turn-taking. Amy tries to find out the reason for Todd's earlier statement about his hair, which arouses her interest enough to direct the question to Todd.

The significance (or likewise the insignificance) of certain participants in the community often surfaces through turn-taking analysis. From the same sample as above (see Excerpt 1, 71-76), all conversation come to a halt when Annie logged on to the session. The members welcome her with all kinds of emoticons to show their excitement in 'seeing' her. On the other hand, no one acknowledges Ted's presence until he starts to join the conversation (see Excerpt 2, 93). His statement regarding Rita's hair does not generate the welcoming responses that Annie receives. It is possible that Ted has been observing the conversation for some time to find the right moment in joining the dialogues or perhaps he has just logged on. Either way, he does not seem to care about any acknowledgment of his presence nor is he concerned with invoking any further actions in announcing his presence.

Sequence of Conversation

Excerpt 2

77.	*amy*	Todd, why will you have small animals in your hair?
78.	*Todd*	Because I have so much of it.....
79.	*Todd*	Even the guy who does my hair jokes about it
80.	*rita2a*	Todd, LOL, I have straight used-to-be-blonde hair.
81.	*Todd*	It's naturally curly, which I now view as a blessing...but it is a chore to take care of
82.	*kevin*	todd you are not a blonde you are definitly a redhead I can tell from here
83.	*Todd*	Mine is blonde..with a little help. I have a lot of gray for someone who is 35, but well covered
84.	*Todd*	kevin...I do have some red in there too LOL
85.	*kevin*	Rita, you are a brunette
86.	*rita2a*	Todd, my Lit Coach (who I'm going with) isn't going to Keene's presentation--so I'll be alone there. I'll definitely look for you.
87.	*kevin*	you guys are ruining my ESIA member schema
88.	*rita2a*	kevin, ROFL
89.	*Todd*	I told my fellow teachers I'm dropping my bag at the hotel and heading over. They can check me in
90.	*AnnieL*	The presentation is next week?
91.	*Todd*	Yes, Annie. I get to meet Rita
92.	*rita2a*	yup Annie
93.	*ted*	Rita is NOT a brunette
94.	*kevin*	I know that now, I'm trying to adjust
95.	*Todd*	ted, are you a brunette?
96.	*AnnieL*	No, Rita is blue.

97.	*Todd*	kevin...and you?
98.	*Terri*	Terri's gray -- just letting it do its thing.
99.	*ted*	natch
100.	*rita2a*	ROFLOL
101.	*kevin*	What do you think Todd?
102.	*Todd*	Hmmmmmmmmmm, ted, blond?
103.	*Todd*	Ummm, well, uh, blonde? kevin?
104.	*Todd*	Thanks everyone for leaving me out here all alone. LOL

Very often, the sequence interactional organization goes hand in hand with turn-taking organization. The basic concept of sequence organization is the implied response that is to follow a prior dialogue. For the reason of relevance to online conversation under discussion, response actions and pre sequence actions were the focus of this analysis. Response actions are frequently displayed in various forms of interactional conversation. Responses such as agreement, disagreement, acceptance, rejection, or correction (Goodwin & Heritage, 1990) are often associated with the promptness in responding to the speaker who initiates the statement, question or comment. Pre-sequence actions are characterized by "preface requests, jokes and stories...news announcement...and invitations" (Goodwin & Heritage, 1990, p. 297). They often cause a change of action or response on the part of the hearer when faced with conflict.

In "Hair Talk", Todd was the first to disclose information about his curly blonde hair (see Excerpt 2, 77, 78, 79, 81, 83, and 84) when Amy asked about the "small animals" in his hair (see Excerpt 2, 77). By having an open-ended sentence in his first reply (see Excerpt 2, 78), Todd seems reluctant to disclose any information about his hair by avoiding to answer Amy's question. However, he seems to change his mind by finishing with his reply on the next two entries of response (see Excerpt 2, 78, 81). From then on, Todd wants to find out the hair color of others by poking for information (see Excerpt 2, 95, 97, and 103). Posting his question to Kevin (see Excerpt 2, 97), Todd generates a seemingly thought provoking responsive action (see Excerpt 2, 101) from Kevin. Instead of answering the question that Todd posts (see Excerpt 2, 97), Kevin raises a question in return, (see Excerpt 2, 101) which leads Todd to ponder on giving the response (see Excerpt 2, 102). Kevin's question (see Excerpt 2, 101) also signifies a pre-announcement that induces Todd to guess rather than to give him the answer. In the same token, the question that Todd directs to Ted regarding his hair color (see Excerpt 2, 95) causes a similar response reaction from Ted (see Excerpt 2, 99). The manipulation of responses by Kevin and Ted enables them to tease Todd. Todd's frustration in getting no reply from others is evident in his response in line 104.

Content of Conversation

Excerpt 3

71.	*Terri*	(((Annie)))	Create social energy
72.	*Todd*	Hi Annie!!!!!!!!	
73.	*rita2a*	(((Annie)))	
74.	*amy*	Hi Annie!	
75.	*kevin*	Hey Annie!!!	
76.	*AnnieL*	Hi Everyone!	
86.	rita*2a*	Todd, my Lit Coach (who I'm going with) isn't going to Keene's presentation--so I'll be alone there. I'll definitely look for you.	Members appear to know each other well
91.	*Todd*	Yes, Annie. I get to meet Rita	
117.	*rita2a*	Todd,I'm a lot older than you, but I'll probably leave the cane at home.	
118.	*kevin*	And I have thin thighs, long legs and buns of steel too	

The content of this seemingly irrelevant “Hair Talk” conversation reflects a distinctive tone and style. Excerpt 3 offers evidence that these participants really know their fellow Book-Talkers in ways that suggest a sense of belonging and true insights into identity. Participants appear to be very comfortable with each other, and they seem to know each other well enough as to tease (perhaps even bait) each other.

Discussion

These seemingly non-sensical online interactions of the "Hair Talk" conversation derive purpose in their potential to support dimensions of activities inherent in communities of practice. Specifically, within the dimension of mutual engagement, members of the community discover ways to interact with each other, build relationships, establish identities, learn about others' identities and explore their strengths and weaknesses. Regardless of the variety of these identities, the participants get to know each other "as presented." More important than confirming a genuine identity is the impulse to know and mutually engage each other. The self-directed and continual greetings when someone logs in reveal the willingness of Book Talk participants to develop and maintain engaging, welcoming mutual relationships (see Excerpt 1, 71-76). Their greetings also signify their acknowledgment of others' identities. For instance, through other-directed organization, Amy, one of the eight participants, learns how to develop a repertoire (styles, actions, artifacts) with Todd. Her initiation in questioning about Todd's hair is intentional and meaningful (see Excerpt 1, 77).

Members in a community tend to share a common goal (joint enterprise). They learn through the process of understanding and negotiating. In the process of working together to solve disagreements and conflicts, members develop mutual accountability toward other members and the enterprise as a whole. Through the engagement with each other, the members also recognize the dynamics of their enterprise and "attempt, neglect, or refuse to make sense and to seek new meanings" (Wenger, 1999, p. 81). From the specimen, it is evident that Todd's frustration indicates his struggles in defining the enterprise and his coping with the conflict. His effort in resolving the confusion at hand indicates his accountability to the other participants. The way that Ted and Kevin lead Todd into a frustrating situation demonstrates their understanding of their own community (albeit put to somewhat negative ends), the members, and the boundaries within the community (see Excerpt 2, 93-104).

According to Rogers (2000), "meaning is negotiated in a community through its shared repertoire. This repertoire refers to the fact that there is a pool of resources that members not only share but also contribute to and therefore renew" (p. 388). New ideas are generally created through the shared repertoire. For instance, despite the fact that this particular online chat is a book discussion session, the participants have no reservation to renegotiate the meaning or purpose of the discussion session (see Excerpt 2, 87-94). They develop their own style in relating to and in learning about each other. These interactions also indicate a strong sense of belonging. They also learn to engender the atmosphere of the chat environment by showing excitement in welcoming others (see Excerpt 3, 71-76), and being playful.

Conclusion

Initially, the selected Book Talk conversation appeared to be trivial and off task, especially considering the fact the participants were reading teachers in a book discussion session. The commentary on each other's hair color does not appear to have any bearings on the rest of the conversation related to upcoming conferences, specific books and upcoming holidays at school. In fact, the conversation seemed somewhat immature. On the other hand, this type of mundane conversation, which is prevalent in many ordinary social conversations and interactions, may possibly have significant implications.

The selected specimen of exchanges were primarily non-informative (Burnett, 2000). Within the 'non-informative' frame, the analysis showed that this kind of seemingly irrelevant online interaction supported a) mutual engagement, b) joint enterprise, and c) development of a shared repertoires - three activities related to the maintenance of coherence in a community of practice (Wenger, 1999).

The relevance of seemingly off topic or tangential utterances should be no surprise to those who are familiar with conversation analysis of talk in face-to-face contexts. However, systematic analysis of how these types of interactions play out in on-line chats has been lacking. These reading teachers log in week after week to Teacher.Net seeking out other teachers. For them it appears that there is potentially much sense in the nonsense of seemingly non-informative online interactions as they may relate to the development of a community of practice. In fact, the insights reported here about these non-informative interactions may help to explain the persistence of these voluntary, online professional communities. Further analysis of the details on online talk-in-interaction may reveal additional information regarding how online talk relates to the building, maintenance or erosion of communities of practice.

References

Agres, C., Igbaria, M., & Edberg, D. (1998). The virtual societies: Forces and issues. *The Information Society, 14*(2), 71-82.

Bott, A. (2002). Partnerships. *Teachers.Net.* Retrieved December 2,2002, from http://teachers.net/partnerships/

Burnett, G. (2000, July). Information exchange in virtual communities: A typology. *Information Research, 5*(4). Retrieved Nov 14, 2002, from http://informationr.net/ir/2005-2004/paper2082.html

Burnett, G., & Bonnici, L. (2003). Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. *Library & Information Science Research, 25*, 333-351.

Cummings, J. N., Butler, B., & Kraut, R. (2002, July). The quality of online social relationships. *Communications of the ACM, 45*(7), 103-108.

Ferry, B., Kiggins, J., Hoban, G., & Lockyer, L. (2000). Using computer-mediated communication to form a knowledge-building community with beginning teachers. *Educational Technology & Society, 3*(3), 10.

Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology, 19*, 283-307.

Gunawardena, C., Lowe, C., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 397-431.

Herring, S. (1999, June). Interactional coherence in CMC. *Journal of Computer-Mediated Communication, 4*(4). Retrieved December 12, 2002, from http://www.ascusc.org/jcmc/vol2004/issue2004/herring.html

Hutchby, I. (2001). *Conversation and technology: From the telephone to the Internet*. Malden, MA: Polity Press/Blackwell Publishers, Inc.

Johnson, C. M. (2001). A survey of current research on online communities of practice. *Internet and Higher Education, 4*, 45-60.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Educational Technology & Society, 5*(1), 8-22.

Lapadat, J. C. (2002, July). Written interaction: A key component in online learning. *Journal of Computer-Mediated Communication, 7*(4). Retrieved May 12, 2003 from http://www.ascusc.org/jcmc/vol2007/issue2004/lapadat.htm.

Mazur, J. (2004). Conversation analysis for educational technologist: Theoretical and methodological issues for researching the structures, processes and meaning of on-line talk. In D. Jonassen (Ed.), *The Handbook of Research for Educational Communications and Technology* (2nd ed.). New York: Simon & Schuster.

Muramatsu, J., & Ackerman, M. S. (1998). Computing, social activity, and entertainment: A field study of a game MUD. *Computer Supported Cooperative Work: The Journal of Collaborative Computing, 7*, 87-122.

Preece, J. (2000). *Online communities: Designing usability, supporting sociability*. New York: John Wiley & Sons.

Rogers, J. (2000). Communities of practice: A framework for fostering coherence in virtual learning communities. *Educational Technology & Society, 3*(3), 384-392.

Schoberth, T., Preece, J., & Heinzl, A. (2002). *Online communities: A longitudinal analysis of communication activities.* Paper presented at the 36th Hawaii International Conference on System Sciences, Hawaii.

Soukup, C. (1999). The gendered interactional patterns of computer-medicated chatrooms: A critical ethnographic study. *The Information Society, 15*, 169-176.

ten Have, P. (2000). *Doing Conversation Analysis: A Practical Guide*. Thousand Oaks, CA: Sage Publications, Inc.

Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet*. New York: Simon & Schuster.

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Effective Problem-Based Instruction: A Comparative Study of Instructional Methods and Student Characteristics

John R. Mergendoller
Buck Institute for Education

Nan L. Maxwell
California State University

Yolanda Bellisimo
College of Marin

Introduction

This study examined the potential differences between Problem Based Learning (PBL) and traditional instructional approaches in building high school students' knowledge of macroeconomic concepts and principles. Using a within-teacher, quasi-experimental design with data from 246 students in 11 classes taught by five teachers, we found a statistically significant ($p < .05$) difference between the problem-based and traditional Lecture/Discussion approach classes in the development of students' economic knowledge, with students in the problem based classes learning more. Results suggest that PBL effectiveness is differentially associated with the following student characteristics: verbal ability, interest in economics, and problem solving efficacy.

Educational reformers seeking to make schools and classrooms more effective learning environments have frequently proposed restructuring traditional curriculum and instruction to engage students in meaningful problem solving (Cognition and Technology Group at Vanderbilt [CTGV], 1997; Hiebert, Carpenter, Fennema, Fuson, Human, Murray, Alwyn, & Wearne, 1996, May). Problem-Based Learning (PBL) is an instructional approach where students are confronted with simulated, real-world problems, and is frequently advanced as a powerful and engaging learning strategy that leads to sustained and transferable learning (e.g., Bransford, Sherwood, Hasselbring, Kinzer, & Williams, 1990; Jones Rasmussen, & Moffitt, 1996; Stepien & Gallagher, 1993; Stepien, Gallagher, & Workman, 1993). By engaging students in a realistic problem that reflects the context and constraints of the "real world," and by requiring students to clarify the problem and to conduct research necessary to solve the problem, it is argued that PBL encourages students to retain newly gained knowledge and solution strategies, fosters the development of self-directed learning strategies, and enables them to apply what they have learned to new and unfamiliar situations (Blumberg, 2000; CTGV, 1997; Maxwell, Bellisimo, & Mergendoller, 2001).

PBL deviates from more conventional instructional strategies by restructuring traditional teacher/student interaction toward active, self-directed learning by the student, rather than didactic, teacher-directed instruction (e.g., Barrows, 1988; Birch, 1986; Savery & Duffy, 1994; Smith & Ragan, 1999; Stepien & Gallagher, 1993; Torp & Sage, 1998). In PBL, teachers coach students with suggestions for further study or inquiry but do not assign predetermined learning activities. Instead, students pursue their own problem solutions by clarifying a problem, posing necessary questions, researching these questions, and producing a product that displays their thinking. These activities are generally conducted in collaborative learning groups, and these groups often solve the same problem in different ways and arrive at different answers.

The design of the PBL instructional approach used in the current study (Maxwell, et. al., 2001) is instantiated in a series of curriculum units focused on the knowledge, concepts, and principles that make up the American high school economics curriculum (Buck Institute for Education, nd). These units, which can take from one day to three weeks to complete, depending upon the unit, scaffold and, to some degree, constrain teacher and student behavior. Each unit contains seven interrelated phases: entry, problem framing, knowledge inventory, problem research and resources, problem twist, problem log, problem exit, and problem debriefing. Student groups generally move through the phases in the order indicated, but may return to a previous phase or linger for a while in a phase as they consider a particularly difficult part of the problem. The teacher takes a facilitative role, answering questions, moving groups along, monitoring and sanctioning positive and negative behavior, and watching for opportunities to direct students to specific resources or provide clarifying explanations. In this version of PBL, students do not learn entirely on their own; teachers still "teach," but the timing and extent of their instructional interventions differ from teachers using traditional Lecture/Discussion approaches. PBL teachers wait for teachable moments when students want to understand specific content or recognize that they must learn something before intervening or providing needed content explanations.

PBL: A Look at the Evidence

Although the theoretical basis for the PBL argument is compelling (Norman and Schmidt, 1992; Regehr & Norman, 1990), little empirical literature exists on the impact of PBL at the high school level. The bulk of the research conducted on PBL instruction has taken place in medical schools (i.e., http://meds.queensu.ca/medicine/pbl/PBLAbstracts.htm), where the PBL instructional model is increasingly at the heart of curriculum reform efforts (Armstrong, 1997; Kaufman, 1985). Reviewers who have examined PBL medical school research have reached contradictory conclusions. Albanese and Mitchell (1993) concluded that problem-based instructional approaches are less effective in teaching basic science content (as measured by Part I of the National Board of Medical Examiners Exam), while Vernon and Blake (1993) reported that PBL approaches were more effective in generating student interest, sustaining motivation, and preparing students for the clinical interactions with patients. Berkson (1993) found that "the graduate of PBL is not distinguishable from his or her traditional counterpart"; this conclusion is consistent with a number of studies have shown no statistically significant difference in learner performance compared to students receiving lecture-based instruction (Albano, Cavallo, Hoogenboom, Magni, Majoor, Manenti, Schwirth, Steigler, & Van, 1996; Blake, Hosokawa, & Riley, 2000; Chang, Cook, Maguire, Skakun, Yakimets, & Warnock, 1995; Farquhar, Haf, & Kotabe, 1986; Kaufman & Mann, 1988; Login, Ransil, Meyer, Truong, Donoff, & McArdle, 1997). Culver (2000) conducted a metanalysis of literature reviews comparing the impact of PBL and Lecture/Discussion instruction and concluded that there is "no convincing evidence that PBL improves knowledge base and clinical performance…." Culver argues that the effects reported in the literature were either too small to be of consequence (generally less than .2 SD), or resulted from selection bias and other methodological defects. In response to Culver, Norman (2001) disputed the general approach of using high-stakes examinations, such as the National Board of Medical Examiners Exam, as a comparative outcome measure. He pointed out that many medical students "cram" or take special preparation courses to prepare for this exam. As a result, the impact of a curriculum design may well make a minor contribution to exam results (Norman, 2001).

Problems abound in generalizing research conducted on students in medical schools to a high school population (Maxwell, et. al., 2001). Medical students are an elite group with superior verbal and quantitative skills. They are older than high school students and their intellectual development has progressed further. They are, presumably, more experienced with and accomplished in the use of hypothetical-deductive reasoning. They have chosen to attend medical school and view their training as instrumental to future occupational success. Given these differences in student characteristics and learning context, it is dubious that findings based on research with medical students can be applied directly to high school courses structured around a PBL format and enrolling a diverse group of students.

Little research has been conducted within high schools comparing the effectiveness of PBL and traditional instructional approaches. Mergendoller, Maxwell, & Bellisimo (2000) compared the learning and attitudes of high school students studying economics using problem-based and lecture discussion methods. They found no statistically significant differences on unit-specific learning outcomes, although there was a difference in changes in general economics knowledge measured at the beginning and end of the semester, with the Lecture/Discussion classes learning more. Visser (2002) compared the effects of problem-based and lecture based instruction on student problem solving and attitudes in a high school genetics class. She found statistically significant differences ($p<.05$) in learning outcomes and motivation for students in the PBL and Lecture/Discussion treatments, with the PBL students reporting less motivation and learning less while recounting more confidence in their learning. Gallagher, Stepien, & Rosenthal (1992) compared the spontaneous problem solving of two groups of gifted high school students: one group had been enrolled in a problem-based Science and Society course, and a comparison group who were not enrolled in the problem-based course. They found that students enrolled in the problem-based course were more proficient in "problem finding" and engaged in problem-solving more successfully and spontaneously than the comparison students (who had not been taught a specific problem-solving process). Given the lack of decisive evidence that a PBL instructional approach is more effective than traditional Lecture/Discussion methodology, we hypothesized that in the current study there would be no difference in learning outcomes between students in PBL and traditional instructional environments.

In addition to incomplete knowledge regarding the effectiveness of PBL instructional approaches with high school students, we know little about the relationship of individual differences *among* high school students to content learning in PBL instructional environments. In a review of the implications of cognitive theory for problem-solving instruction, Fredericksen (1984) noted, "there is considerable evidence that aptitude-treatment interactions exist." (Aptitude-treatment interactions occur when certain treatments have differential effects on students with different aptitudes.) We are interested in two general categories of aptitudes which may shed light on the efficacy of PBL environments with different students.

The first category of aptitudes include academic ability and subject matter interest. These relatively stable student characteristics are of interest as some authors have argued that lower ability and chronically uninterested students, who often do not thrive in traditional, Lecture/Discussion learning situations, are more likely to succeed in content rich, socially collaborative, contextually meaningful learning environments, such as those established in well-implemented PBL (e.g., Delisle, 1997; Glasgow, 1977; Jones et al., 1996). Our review of the PBL research literature, however, revealed no empirical studies suggesting that PBL is an effective instructional approach for lower ability high school students. In fact, it may be just the opposite. One of the best known American high schools incorporating a PBL approach is the Illinois Mathematics and Science Academy (www.imsa.edu). In addition to advocating the use of PBL instructional methodology, IMSA conducts research on the impact of PBL and trains teachers from other schools in PBL methodology. IMSA students, however, are chosen in a highly selective admission process and demonstrate superior ability in mathematics and science. A previous study by the current authors (Mergendoller, et al, 2000) found that verbal ability was positively associated with successful learning in *both* PBL and traditional, Lecture/Discussion high school courses. Given the scant research on problem-based instruction in high school, it is evident that more research is needed before claims of PBL's superior efficacy with lower-achieving students can be accepted.

In addition to academic ability, there is a question of whether interest in the subject matter being addressed is related to attainment in PBL learning environments, as students who are interested in learning a particular subject may be more willing to engage in the complex cognitive and interactional tasks required by PBL. Such active intellectual and social engagement is generally more demanding than listening to a lecture or participating in a class discussion (Blumenfeld, Mergendoller, & Swarthout, 1987; Doyle, 1983). Throughout a PBL experience, students take an active role in their learning as they discuss and decide on problem-solving strategies, divide research and other responsibilities among group members, communicate the results of their research back to the group, and finally craft a problem solution which is often presented to an external audience. In response to these considerations, we hypothesized that verbal ability and student content interest would be related to learning in PBL environments, and students with superior verbal ability and stronger interest in learning economics would learn more in PBL classes.

A second category of student aptitudes include those which are more directly related to the task and interactional demands of the PBL learning environment. Meyer, Turner, & Spencer (1997) reported that individual differences in motivation and self perception influenced mathematics attainment in investigative, activity-based group learning, an instructional modality with many characteristics in common with PBL. Ethnographic research by Anderson, Holland, & Palincsar (1997) documented how interpersonal dynamics and perceptions of the capability of other group members can alter the task demands and participatory behavior, and can limit the learning opportunities available to less academically talented group members.

Given this research and our own observations of PBL learning environments, we hypothesized that students who preferred to learn in groups and who perceived themselves to be competent problem-solvers would learn more in PBL learning environments.

Finally, Brown, Johnson, Mayall, Boyer, Reis, Butler, Weir, & Florea (2002) examined whether gender is associated with differences in PBL participation and learning, and identified little differentiation between males and females. We hypothesized that in the current study there would be no gender differences in relation to learning outcomes.

The following study addresses three questions:

1) Is the Problem Based instructional method more effective than a traditional Lecture/Discussion approach in teaching high school students about macroeconomic concepts?
2) Is verbal ability or content interest associated with student success in PBL learning environments?
3) Is student preference for group work, and perceived problem-solving efficacy associated with student success in PBL learning environments?

Method

Our study employed a within-teacher, quasi-experimental design with non-random assignment of students to classes. Five veteran teachers at four different high schools participated in the study. All of the high schools were located in a large metropolitan area in Northern California. Two of the high schools were suburban, and two were urban. To control for teacher effects, all teachers taught the same macroeconomics content using a PBL approach with one or more classes and a traditional Lecture/Discussion approach with one class. Teachers were allowed to select which class they would instruct using a Lecture/Discussion approach, but this choice was made before the school year began, and before teachers had received their class lists. Consequently, teachers had no advance indication of the student composition of each class. PBL and traditional classes were distributed throughout the school day, with four of the five teachers teaching the PBL and traditional classes within 2 periods of each other. The remaining teacher's PBL and traditional classes were within 3 periods of each other.

The focus of the units in both the PBL and Lecture/Discussion classes consisted of the macroeconomics content defined by the *National Voluntary Content Standards in Economics* (www.economicsamerica.org), and *A Framework for Teaching Basic Economic Concepts* (Sanders and Gilliard, 1995).[1] The problem-based unit, *The President's Dilemma*, casts students as teams of economic advisors to the president during a time when the increasing cost of oil has resulted in sluggish economic growth, high unemployment and high inflation. Solution of this problem requires students to recommend fiscal and monetary policy alternatives that will address these economic problems and get the economy growing again. To determine the best policy alternatives, students must develop a knowledge of monetary and fiscal policies, gross domestic product, unemployment and inflation, economic incentives, public policy alternatives and costs. As the problem unfolds, students discover that scarcity dictates societal tradeoffs and opportunity costs in pursuing a healthy economy.

This problem is ill-structured in that information necessary to solve the problem is not "pre-packaged," but exists in a variety of places. Students' judgments of relevant and irrelevant information and their definition of the problem being solved changes as they delve deeper into the problem. There are also, as in real-world problems, multiple correct solutions to the problem as well as multiple incorrect ones (Maxwell et al., 2001). The problem, although allowing for student discovery and independent learning, proceeds in a structured manner. Students work in groups, they clarify the nature of the problem and determine what economic concepts and relationships are necessary to solve it, and they undertake the research and reading necessary to understand the relevant economic theory. The problem concludes with a presentation of the solution each group has fashioned to an audience of interest group representatives (e.g., the elderly, labor unions, business owners, etc.). These representatives (usually played by other teachers or interested parents) are primed with specific questions that elicit students understanding – and misunderstanding – of economic concepts and principles (e.g., "Given the fiscal policy actions you have proposed, what would be the impact if the Federal Reserve unexpectedly raised the discount rate?"). Although one student gives the group's speech, questions are addressed to individual group members. This procedure, and the potential for public embarrassment, increases the pressure on students to understand the economic concepts at the heart of the unit.

Teachers were asked to spend the same amount of time and address the same concepts in both the Lecture/Discussion and PBL classes. All teachers had attended weeklong training workshops (under the guidance of a university economics professor and co-developer of the problem) to prepare them to use the PBL economics units in their classes. Two of the five participating teachers have worked as trainers for subsequent workshops. All instructional resources necessary to teach the PBL units were provided, including a carefully prepared curriculum guide and tips and strategies for guiding students through the problems. Conversations with teachers as they taught the units and at debriefings when they had completed the unit suggested that the PBL and Lecture/Discussion approaches were implemented as intended by the materials developers and researchers.[2]

A total of 346 twelfth-grade students in 11 classes completed one or more of the instruments used in the study. The following data analysis is based on data collected from the 246 students who completed the pre- and post-macroeconomics knowledge instrument and the verbal ability measure. These students make up 71% of all the students in the classes. The amount of student attrition is testament to the elevated absence rates among graduating seniors during the second semester of the senior year.

At the beginning of the semester, students in both the traditional and PBL classes completed the aptitude measures (academic ability, attitude toward economics, preference for group work and problem-solving efficacy). Immediately before (pretest) and immediately after (posttest) the macroeconomics unit, students completed a multiple choice content test.

Academic ability was measured using the Quick Word Test: Level 1 (Borgatta, 1964). Each item consisted of a target word in capital letters followed by four lower-case words. Students were asked to circle the appropriate synonym for the target word. A student's score was calculated by summing the correct answers. The test authors report strong validity and reliability, including correlations greater than 0.80 with the Verbal, Total, and IQ scales of the Weschler Adult Intelligence Scale and split-half reliability coefficients of greater than 0.90 (Borgatta & Corsini, 1964, 1967).

We searched for an instrument appropriate to measure high school students' interest in learning economics, and did not find anything suitable. The instruments we reviewed assumed a basic knowledge of economics, and contained items such as "I enjoy economics" or "Economics is practical" (Hodgin, 1984). Since the majority of high school students have never studied economics, and consequently have incomplete or erroneous knowledge of economic concepts and principles, asking them about interest in economics is like asking them their interest in biophysics – they may have heard the word, but generally don't know enough about the concept to express a valid opinion. A result we designed our own instrument asking students about their interest in learning about economic issues. The instrument consisted of the stem: "How interested are you in reading newspaper and/or magazine articles about. . . " followed by four items describing the economic plight of various groups (e.g., economic issues faced by the poor) and two items describing general economic issues (e.g., unemployment). Students responded on a five-point Likert scale running from Very Interested to Not Interested. We calculated scores by taking the mean response across all six items. Cronbach's alpha for the instrument was 0.80.

We measured preference for group work using four items sharing the common stem, "When I work with my classmates in small groups, I usually find that . . . ." Items included: "it does not help me learn, " "it gives me new ways to think about what we are studying," and "it is an excellent way to study for tests." Students indicated their response on a five-point Likert scale running from Strongly Agree to Strongly Disagree. After reversing negatively worded items, we calculated student scores using the mean of the four items. Cronbach's alpha for this scale was 0.79.

We measured students' perceptions of their own problem solving skills by asking them about the behaviors required in PBL learning environments. The instrument consisted of the stem: "I have difficulty solving problems when . . . . " Items described processes of problem solving, negotiation, and discussion such as "I have to find my own resources and information," and "I have to argue my own point of view. Students responded on a five-point Likert scale running from Strongly Agree to Strongly Disagree, and their scores reflect the mean of the six items. Cronbach's alpha for this scale was 0.82.

We created a unit-specific content test using 16 four-part, multiple choice items drawn from the Test of Economic Literacy (Soper & Walstad, 1987) and the test bank accompanying a widely used high school economics textbook (Marlin, Mings, & Swanson, 1995). The items addressed the full range of cognitive objectives (knowledge, comprehension, application, analysis, and evaluation) described by Bloom, Englehart, Furst, Hill, & Krathwohl (1956), and were focused on the specific concepts to be covered in the classes. Students' scores were obtained by summing the number of correct items. Inspection of histograms for both the pretests and posttests suggest a normal distribution with no outliers.

Results

Table 1 displays descriptive information about the students and classes participating in the research, as well as mean posttest-pretest gain and effect size by teacher and instructional condition. As can be seen, there was considerable heterogeneity among teachers in the verbal ability of their students, ranging from a mean of 36.67 for students in Teacher E's PBL class to a mean of 59.51 for students Teacher C's Lecture discussion class. Similar heterogeneity is seen in mean pretest score, ranging from 4.42 to 8.57. For four of the teachers, comparisons of the verbal ability and pretest scores among students in the PBL and traditional instruction approaches did not reach statistical significance, suggesting that for each of these teachers the students in the PBL and traditional classes were of similar academic ability. For one teacher (D), there was a statistically significant difference between the pretest scores, but not the verbal ability scores, of students in the PBL and traditional instructional approaches. This teacher also showed the greatest difference between the gains demonstrated by students in the PBL and traditional classes.

Independent-Samples T Tests were used to examine whether students in the PBL and Lecture/Discussion classes showed statistically significant differences in their Verbal Ability, Interest in Learning Economics, Problem Solving Efficacy, and Preference for Group Work. Across instructional conditions, there were no statistically significant differences, but Interest in Learning Economics came close ($t = 1.89$, $p = .06$) with PBL students reporting more interest at the beginning of the study than Lecture/Discussion students.

Independent-samples T tests were also used to compare posttest-pretest change scores by instructional condition across all teachers and within teachers. Across all classes, the posttest/pretest gain was +1.48 (SD=2.52) for the PBL students and +.82 (SD=2.81) for the Lecture/Discussion students. This difference is statistically significant, $t = 1.94$, $p = .05$, and equivalent to an effect size of .59 for students in the PBL instructional approach and .29 for students in the Lecture/Discussion approach. This indicates that on average students in the PBL classes gained more on the macroeconomics content test than students in the traditional classes. Looking at the within-teacher comparisons displayed on Table 1, we see that PBL classes taught by four of the five teachers gained more than the traditional classes, although only two of these comparisons reach statistical significance. For one teacher (C), students gained more in the traditional class, although this difference was not statistically significant at the .05 level.

The above data suggest that the answer to the first research question is that the PBL instructional approach was more effective than a traditional Lecture/Discussion approach in helping students to learn basic macroeconomic concepts, contrary to what we had hypothesized based on the medical education literature.

The remaining research questions examine whether the PBL approach is more effective for students with certain characteristics. For this analysis, we treated each student characteristic separately and created
tertiles containing students with "high," "medium," and low levels of each characteristic (except for gender).

Table 1. *Means and Standard Deviations of Verbal Ability, Pretest, and Posttest-Pretest Change by Teacher and Instructional Condition*

Teacher		N Students	Verbal Ability	Pretest	Posttest-Pretest Change	Effect Size
A	**PBL**	44 (2 classes)	56.61 (12.18)	7.68 (2.61)	1.36 (2.18)	.50
	Lecture/Discussion	19	56.78 (17.35)	7.21 (3.28)	-.21 (2.96)	-.06
	Difference		-.17	.55	***1.57****	.61
B	**PBL**	26	42.46 (11.47)	4.42 (1.78)	1.42 (2.86)	.70
	Lecture/Discussion	24	41.70 (11.45)	5.13 (2.69)	1.04 (2.69)	.35
	Difference		.77	-.70	.38	.14
C	**PBL**	23	58.39 (18.98)	7.78 (2.61)	1.09 (2.39)	.44
	Lecture/Discussion	23	59.51 (12.51)	8.57 (2.48)	2.35 (3.02)	.95
	Difference		-1.12	-.78	-1.26	-.46

D	**PBL**	21	49.76	5.10	2.43	.93
			(14.91)	(2.61)	(2.40)	
	Lecture/Discussion	21	52.29	7.14	-.19	-.06
			(14.21)	(2.78)	(2.62)	
	Difference		-2.52	***-2.05****	***2.62*****	1.04
E	**PBL**	25	36.67	5.88	1.32	.57
			(11.60)	(2.33)	(2.88)	
	Lecture/Discussion	20	40.92	5.10	.85	.38
			(14.33)	(2.13)	(2.03)	
	Difference		-4.25	.78	.47	.19
ALL TEACHERS						
	PBL	139	49.64	6.37	1.48	.59
			(15.86)	(2.77)	(2.52)	
	Lecture/Discussion	107	50.14	6.63	.82	.29
			(15.72)	(2.97)	(2.81)	
	Difference		-.50	-.25	***.66****	.25

NOTE: * = p < .05, ** = p < .01

Thus the "high" tertile included students whose scored in the 67th to 99th percentile for that characteristic. Conversely, the "low" tertile included students scoring in the 0 to 33rd percentile. The "medium" tertile contained the remaining students. After splitting the population into these three groups, we conducted four separate Analyses of Variance within each instructional condition. Post-hoc comparisons with Bonferroni corrections were used to evaluate whether the posttest-pretest score differed between students in any of the tertiles. We found no differences statistically significant at the .05 level within either the PBL or Lecture/Discussion students.

For each variable, we also conducted Independent-Samples T-Tests within tertile comparing mean posttest-pretest scores of students in the PBL and Lecture/Discussion classes. Table 2 displays the data used in

Table 2. *Mean Pretest and Posttest-Pretest Change within Instructional Condition by Verbal Ability, Interest in Learning Economics, Problem Solving Efficacy, Preference for Group Work Tertiles*

Variable	Tertile	Condition	N	Pretest	Posttest-Pretest Change	SD	*t*	Effect Size
Verbal Ability								
	High						.24	.05
		PBL	44	8.16	1.45	2.25		
		L/D	39	8.38	1.31	3.30		
	Medium						1.67	.41
		PBL	43	6.12	1.84	2.54		
		L/D	24	5.83	.67	3.10		
	Low						1.66	.40
		PBL	49	4.92	1.22	2.71		
		L/D	31	5.45	.26	2.08		
Interest in Learning Economics								
	High						***2.21****	.50
		PBL	55	6.56	1.24	2.68		
		L/D	31	7.39	-.10	2.70		
	Medium						.92	.21
		PBL	43	6.56	1.67	2.59		

		L/D	34	5.94	1.12	2.68		
	Low						.61	.14
		PBL	41	5.93	1.61	2.22		
		L/D	42	6.62	1.26	2.90		
Problem Solving Efficacy								
	High						1.73	.88
		PBL	52	7.62	1.48	2.33		
		L/D	33	7.48	-.58	2.37		
	Medium						.05	-.01
		PBL	38	6.13	1.37	2.47		
		L/D	40	6.18	1.40	2.79		
	Low						1.81	.40
		PBL	49	5.24	1.57	2.78		
		L/D	34	6.32	.38	3.17		
Preference for Group Work								
	High						1.29	.32
		PBL	48	5.73	1.10	2.60		
		L/D	28	6.43	.36	2.09		
	Medium						.84	.17
		PBL	47	6.57	1.53	2.23		
		L/D	44	6.07	1.14	2.26		
	Low						1.42	.32
		PBL	44	6.86	1.84	2.70		
		L/D	35	7.49	.80	3.80		

NOTE: * = p < .05

this analysis, the T-Tests results, and the effect size for each comparison. We include effect size information as many researchers believe that effect sizes provide important additional information about the magnitude of differences, whether or not they reach statistical significance, which is directly related to sample size (e.g., Huberty, & Pike, 1999; Kier, 1999).

Except for students in the "high" Interest in Learning Economics tertile, there were no statistically significant differences at the .05 level in the mean posttest-pretest change of students in PBL and Lecture/Discussion classes. At the same time there are some interesting differences in effect size to which we will return in the Discussion section.

Discussion

While not wanting to overstate the import of a single study, our results provide some support for those who advocate PBL instructional approaches. Students' content learning in high school economics classes, as measured by a traditional multiple-choice measure, was greater in PBL classes than in Lecture/Discussion classes. We believe this to be a compelling finding, given that the statistically significant difference in posttest-pretest score across all classes is mirrored by differences at the teacher level favoring PBL classes for four out of the five teachers. Across all teachers, the average effect size difference for PBL-Lecture/Discussion comparisons was .25 or 1/4th of a standard deviation. Interestingly, this is roughly the effect size difference reported by Culver (2000) in his metanalysis of the comparative impact of PBL and traditional instruction in medical schools. Unlike Culver, however, we do not consider the size of this difference to be negligible. Instead, we would apply the convention established by Cohen (1988) defining effect sizes of this magnitude to be "small" but not meaningless. Moreover, most students would not consider trivial the mean difference in posttest-pretest score between the PBL and Lecture/Discussion classes. Across all teachers, PBL classes gained .66 more than the Lecture/Discussion classes. This is equivalent to a raw score difference of 4% – or the distance between a B and a B+ in a grading system based on a maximum score of 100%.

We find these results exciting, but at the same time, they leave unanswered many important questions. A key limitation of the current study is the lack of in-depth information about what, exactly, teachers were doing in the PBL classes that distinguished them from the Lecture/Discussion classes, and how these differences were associated with increased student learning. Future research should address this lacunae with observational studies of PBL instructional environments. We believe it of special importance to develop operational concepts that document the processes of problem based learning, and distinguish its essential components. We also believe it important to develop measurement strategies that can be used to assess non-content related outcomes theoretically associated with problem based instructional approaches and espoused by PBL advocates. We are sympathetic to Norman's (2001) argument that it is the test preparation activities engaged in by individual students – rather than the instructional approach used by the teacher – that best accounts for differences in performance on standardized, content-based tests. Well-designed research comparing the impact of PBL and traditional instruction on students' self-management skills and practices and ability to apply a problem-solving algorithm to non school-centered problems is sorely needed. We shall be turning our own research efforts in this direction.

Our next two research questions focused on whether PBL was a more effective learning environment for students with certain characteristics. Here the results displayed on Table 2 are more equivocal. While comparisons of posttest-pretest change by variable tertile within instructional condition were not statistically significant at the .05 level, the effect size differences for PBL and Lecture/Discussion students in different variable tertiles are provocative. We argue that these comparisons provide some evidence that students with different characteristics perform differently within PBL and Lecture/Discussion classes. Consider, for example, the difference in effect size for students in the high (.05), medium (.41) and low (.40) verbal ability tertiles. While there was no meaningful learning difference by instructional condition for the most verbally proficient students, students whose verbal ability was mid-range and below learned more in the PBL classes than they did in the Lecture/Discussion classes. This result can not be accounted for as an instance of "regression to the mean," as medium tertile students in the PBL classes scored slightly higher than the medium Lecture/Discussion students on the Pretest, while the relative ranking was reversed for the low tertile students. In each case, the effect size difference favoring the learning of the PBL students was approximately .40, a small, but not insignificant difference, equivalent to a raw score difference of 6-7%, or the distance between a D+ and a solid C.

Although this is hardly a ringing endorsement of the use of PBL approaches with lower-achieving students, it does, we believe, provide the first empirical evidence – rather than theoretical argument – supporting the efficacy of PBL instructional methodology for students with limited verbal skills, a key component of cognitive ability measures, and consequently, a predictor of school success (Gage, & Berliner, 1997, p. 58 – 59). Further empirical examination of the efficacy of PBL with students who typically do not succeed in school is another important avenue of future research, and we would urge that students' "at-risk" status be ascertained by multiple measures, not just verbal ability.

Instructional approach also appear to affect students differently according to their interest in learning economics. Lecture/Discussion students most interested in learning economics showed little change in mean content knowledge (-.10) between the pre-test and posttest. On the other hand, PBL students with the same level of interest in learning economics gained in content knowledge (+1.24). Although this difference is not statistically significant, it is equivalent to ½ standard deviation – a "medium" effect size according to Cohen's (1988) convention, and equivalent to a raw score gain of 8%. It is tantalizing to argue that students' with more interest in learning economics were able to capitalize on this interest to expand their personal explorations of economics in the PBL classrooms, an activity that could not occur as easily (if at all) in the Lecture/Discussion classrooms.

The effect size differences for problem solving efficacy present a curvilinear (u-shaped) profile and suggest tell another story. Whether a student was in a PBL or Lecture/Discussion classroom did not appear to make a difference for middle tertile students. On the other hand, students in the top and bottom tertile of problem solving efficacy learned more in the PBL classrooms, with the effect size difference between top tertile Lecture/Discussion and PBL students exceeding ¾ standard deviation (+.88). Again, given that this is a "black box" study with no record of student interactions, one can only speculate why this might be the case. We present the following hypothesis as a plausible explanation in hopes that it might suggest a fruitful area for future research.

Published accounts of student interaction in problem-solving groups (e.g., Anderson, et. al., 1997), as well as our own observations during the development of the PBE units suggest that group members vary considerably in the degree to which they take a leadership role. Some group members plunge in and lead the problem solving effort. Others hang back and look to others to assign tasks and monitor results. All teachers who place their students in groups confront "freeloading," where one or two students do the majority of the work for the others. We had this (as well as other) group management problems in mind when we designed the PBE units, and followed Slavin's dictim that maximum group learning occurs when there is individual accountability (e.g., Slavin, 1990). We therefore structured each unit to include two types of individual accountability – an individually administered multiple choice test and a procedure by which all group members are held individually accountable for justifying their problem solution and explaining their understanding the key economic concepts. For the President's Dilemma, this procedure requires group members to explain, individually, the logic behind their economic prescriptions to an audience of interest group representatives such as the elderly and union members. We believe holding students individually accountable for their learning has a definite influence on the nature of the group interaction, and that students who are not confident in solving the problem by themselves reach out to other students for clarification and enlightenment during group research and discussion.

The review of group processes in the classroom by Webb & Palincsar (1996) identifies two individual actions associated with increased learning; 1) giving elaborated explanations to other group members, and 2) applying explanations (either received or self-generated) to solve problems or perform tasks (p. 854). We hypothesize that the PBL students who were confident in their problem-solving ability would be the one most likely to explain and clarify economic ideas for other group members. Similar opportunities for students to clarify other students' economic understandings would not be available in Lecture/Discussion classes. At the same time, students who felt less confident in solving the economic problems by themselves could solicit help from other students and digest and apply economic explanations as they worked through the problem. Once more, similar opportunities might not be available in the Lecture/Discussion classrooms. This analysis is highly speculative, but it does point the way to important future areas of study.

The final student characteristics that merit discussion are Preference for Group Work and Gender. Here PBL-Lecture/Discussion effect size differences by Preference for Group Work tertile are too small (.15) to be meaningful. We suspect that the impact of students' preferences for a certain classroom instructional approach is outweighed by teachers' accountability systems, and the nature of the interaction that occurs in the classroom. In the abstract, students may prefer working by themselves or with others, but once they are actually in Mrs. Jones class their learning is more influenced by environmental and structural factors than their own learning group work preferences. Similarly, we did not find noteworthy differences in "PBL benefit" between female and male students. The difference in PBL – Lecture/Discussion effect size comparisons is relatively similar for both genders, reflecting a raw score difference between of 4 – 5% for each gender.

Looking ahead to further research on PBL we urge that the scope of the research endeavor be expanded. The current study examined student learning within a single, two-week unit. If problem-based instruction is to help students develop the deep, applicable knowledge and analysis skills that facilitate economic literacy, it is likely that students will need to solve multiple problems over the course of a semester or school year. Research should focus on the additive impact of multiple units, and comparisons should be made of PBL – Lecture/Discussion learning gains during initial units when students are first learning how to take advantage of the PBL approach and again when they are familiar with the working of PBL and ready to exploit the learning opportunities it offers.

In closing, we wish to point out that while PBL was more effective than Lecture/Discussion teaching in increasing academic achievement, the size of this increase, although statistically significant, was not great. PBL is not the silver bullet that will drastically increase the achievement of all students. At the same time, our data demonstrate that it did enable the majority of students in our study to learn more. This finding should be liberating to teachers, instructional designers, and researchers who seek alternatives to traditional "sage-on-the-stage" pedagogy. In this study, PBL not only "did no harm," it did some good, and this should encourage educators to tinker with PBL to better understand the classroom conditions and social arrangements necessary to maximize its effectiveness.

This future research and development effort should include a focus on the hard to measure, learning outcomes of sustained content retention and application as well as self-management and problem-solving skills. Inventive methods need to be developed to compare whether students in the PBL classes can *apply* economic knowledge gained in the classroom to real-world situations. This is the critical test for problem-based teaching (Mayer, & Wittrock, 1996; CTGV, 1997), and the outcome that will validate its promise. Further research on this question, as well as more analysis of the optimal configuration of Problem-based teaching are needed.

References

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, *68*, 52 - 81.Albano, M. G., Cavallo, F., Hoogenboom, R., Magni, F., Majoor, G., Manenti, F., Schuwirth, L., Stiegler, I., and van, d., V (1996). An international comparison of knowledge levels of medical students: the Maastricht Progress Test. *Medical Education*, 30, 239-245.

Anderson, C. W., Holland, J. W., & Palincsar, A. S. (1997, March). Canonical and sociocultural approaches to research and reform in science education: The story of Juan and his group. *Elementary School Journal, 97*(4), 359 – 384.

Armstrong, E. G. (1997). *The challenge of problem based learning (2nd ed.).* London: Kogan Page.

Barrows, H. S. (1988). *The tutorial process*, Springfield Ill.: Southern Illinois University School of Medicine.

Berkson. 1993. Problem-based learning: Have the expectations been met? Invited Review, *Academic Medicine*, 10, S79 – 88.

Birch, W. (1986). Towards a model for problem-based learning. *Studies in Higher Education,* 11: 73-82.

Blake, R. L., Hosokawa, M. C., & Riley, S. L. (2000). Student performances on Step 1 and Step 2 of the United States Medical Licensing Examination following implementation of a problem-based learning curriculum. *Academic Medicine*, *75*, 66 – 70.

Bloom, B. S., Englehart, M. B., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives. *The classification of educational goals. Handbook I: Cognitive domain*. New York: Longmans Green.

Blumberg, P. (2000). Evaluating the evidence that problem-based learners are self-directed learners: A review of the Literature. *In* D. H. Evensen & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions*. Mahwah, NH: Lawrence Earlbaum Associates.

Blumenfeld, P. C., Mergendoller, J. R., & Swarthout, D. W. (1987). Tasks as heuristics for understanding student learning and motivation. Journal of Curriculum Studies, 19(2), 135-148.

Borgatta, E.G. (1964). *Quick word test: Level 1-form cm*. New York: Harcourt, Brace and World.

Borgatta, E.G., & Corsini, R.J. (1967). *Quick word test: Supplementary report*. New York: Harcourt, Brace and World, Inc.

Borgatta, E.G., & Corsini, R.J. (1964). *Quick word test: Directions for administering, scoring, and interpreting*. New York: Harcourt, Brace and World, Inc.

Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C. K., & Williams, S. M. (1990). Anchored Instruction: Why we need it and how technology can help. In D. Nix & R. Spiro (Eds.), *Advances in computer-video technology*. Hillsdale, NJ: Erlbaum.

Brown, S. W., Johnson, P. R., Mayall, H., Boyer, M., Reis, S., Butler, M., Weir, K., & Florea, N. (2002). Gender differences in a problem-based learning environment: Do they exist? Paper presented at the 2002 American Educational Research Association Annual Meeting, April 4, 2002, New Orleans, LA.

Buck Institute for Education. (nd). *Problem Based Economics*. Novato, CA: Author.

Chang, G., Cook, D., Maguire, T., Skakun, E., Yakimets, W. W., & Warnock, G. L. (1995). Problem-based learning: its role in undergraduate surgical education. *Canadian Journal of Surgery*, *38*, 13 – 21.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project: lessons in curriculum, instruction, assessment, and professional development.* Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum

Culver, J. A. (2000, March). Effectiveness of Problem-based learning curricula: Research and theory. *Academic Medicine, 75*(3), 259-266

Delisle, R. (1997). *How to use* problem-*based learning in the classroom.* Alexandria, VA: Association for Supervision and Curriculum Development.

Doyle, W. (1983). Academic Work. *Review of Educational Research, 53,* 159 – 199.

Farquhar, L. J., Haf, J., & Kotabe, K. (1986). Effect of two preclinical curricula on NBME Part I examination performance. *Journal of Medical Education*, *61*, 368 – 373.

Frederiksen, N. (1984). Implications of cognitive theory for instruction in problem solving. *Review of Educational Research, 54*(3), 363 - 407.

Gage, N. L., & Berliner, D. C. (1997). *Educational Psychology 6th Edition.* Boston: Houghton Mifflin Company.

Gallagher, S., Stepien, W., & Rosenthal, H. (1992). The effects of problem-based learning on problem solving. *Gifted Child Quarterly, 36*(4), 195 – 200.

Glasgow, N., (1997). *New curriculum for new times: A guide to student-centered, problem-based learning.* Thousand Oaks, CA: Corwin Press, Inc.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Alwyn, O., & Wearne, D. (1996, May). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher, 25*(4), 12-21.

Hodgin, R. F. (1984). *Attitude assessment for research in economics education.* ERIC Document Service ED 248 779.

Huberty, C. J., & Pike, C. J. (1999). On some history regarding statistical testing. In B. Thompson (Ed.), *Advances in Social Science Methodology (Vol. 5).* Stamford, CT: JAI Press, 1 – 22.

Jones, B. F., Rasmussen, C. M., & Moffitt, M. C. (1996). *Real-life problem solving: A collaborative approach to interdisciplinary learning.* Washington, DC: American Psychological Association.

Kaufman, A. (Ed.) (1985). *Implementing problem-based medical education.* New York: Springer.

Kaufman, D. M., & Mann, K. V. (1998). Comparing achievement on the Medical Council of Canada Qualifying Examination Part I of students in conventional and problem-based learning curricula. *Academic Medicine*, *73*, 1211 – 1213.

Kier, F. J. (1999). Effect size measures: What they are and how to compute them. In B. Thompson (Ed.), *Advances in Social Science Methodology (Vol. 5).* Stamford, CT: JAI Press, 87 – 100.

Login, G. R., Ransil, B. J., Meyer, M., Truong, N. T., Donoff, R. B., & McArdle, P. J. (1997). Assessment of precinlical problem-based learning versus lecture-based learning. *Journal of Dentistry Education*, *61*, 473 – 479.

Marlin, Matthew, Mings, T., & Swanson, D. (1995). *Teaching and testing from The Study of Economics: Principles, Concepts and Applications* (5th edition). Guilford, Connecticut: Dushkin Publishing Group/Brown and Benchmark Publishers.

Maxwell, N. L., Bellisimo, Y., & Mergendoller, J. (2001, March/April). Problem-based learning: Modifying the medical school model for teaching high school economics. *The Social Studies, 92*(2), 73-78.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology*. New York: Macmillan, 47 – 62.

Mergendoller, J., R., Maxwell, N. L., & Bellisimo, Y. (2000). Comparing problem-based learning and traditional instruction in high school economics. *Journal of Educational Research, 93*(6), 374-382.

Meyer, D. K., Turner, J. C., & Spencer, C. A. (1997). Challenge in a mathematics classroom: Students' motivation and strategies in project-based learning. *Elementary School Journal, 97*(5), 501 - 521.

Norman, G., R. (2001, April). The Cognitive perspective on learning: Problem-based learning as a facilitator of knowledge transfer. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Norman, G. R., & Schmidt, H. G. (1992). The Psychological basis of problem-based learning: A Review of the evidence. *Academic Medicine, 67,* 557 - 565.

Regehr, G., & Norman, G. R. (1990). Issues in cognitive psychology: Implications for professional education. *Academic Medicine, 71*, 988 - 1001.

Sanders, P. , & Gilliard, J. eds. (1995). *A framework for teaching basic economic concepts: With scope and sequence guidelines, K-12.* New York: Joint Council on Economic Education.

Savery, J. R., & Duffy, T. M. (1994). Problem-based learning: An instructional model and its constructivist framework. *Educational Technology*, *35*(5), 31 – 38.

Slavin, R. E. (1990). *Cooperative Leaning: Theory, research and practice*. Englewood Cliffs, NH: Prentice Hall.

Smith, P., & Ragan, T. (1999). Instructional Design, 2nd Edition. Upper Saddle River, NJ: Merrill Publishing.

Soper, J. C., & Walstad, W. B. (1987). *Test of economic literacy (2nd edition)*. New York: Joint Council on Economic Education.

Stepien, W., & Gallagher, S. (1993). Problem-based learning: As authentic as it gets. *Educational*

Leadership, 50(7), 25- 28.
Stepien, W., Gallagher, S., & Workman, D. (1993). Problem-based learning for traditional and interdisciplinary classrooms. *Journal for the Education of the Gifted*, 16 (d4), 338-357.
Torp, L., & S. Sage. (1998). *Problems as possibilities: Problem-based learning for K-12 education*. Alexandria, VA: Association for Supervision and Curriculum Development.
Vernon, D. T. A., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine, 68*, 550 - 563.
Visser, Y. L. (April, 2002). Effects of problem-based and lecture-based instructional strategies on problem solving performance and learner attitudes in a high school genetics class. Paper presented at the 2002 Annual Meeting of the American Educational Research Association, New Orleans.
Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology*. New York: Macmillan, 841 – 876.

Notes

1. This curriculum was developed by a partnership between high school teachers, an educational research institution, and economics faculty at a university (Bellisimo et al., 1998). The *President's Dilemma* unit is part of a eight-unit PBL economics' curriculum designed for a semester-long high school course, although each of the eight units can be used in isolation. All of the units focus on the core economic concepts of scarcity, opportunity costs, and tradeoffs, as well as unit-specific concepts (e.g., Fiscal Policy, Monetary Policy, Inflation, etc.)
2. More extensive information on the Problem Based Economics units can be found at http://www.bie.org/pbss/pbe/index.php

Implementing Online High School Programs: An Evaluation Study

Steven C. Mills
University of Kansas

M.D. Roblyer
University of Maryland University College

Abstract

Because of its rapid expansion in enrollment and course offerings, a virtual high school undertook a comprehensive evaluation initiative to ensure the viability of its programs and services. This study of over 2600 online student enrollments determined that online delivery of high school courses fulfills a real and practical need in the high school curriculum. This study also determined that achievement in online courses was equal to or better than achievement in regular high school courses.

Introduction

The tremendous growth and expansion of the Internet was an important development in the advancement of delivery methods for distance education courses and in changing the definition of distance education. In the past, distance education was comprised of courses where the instructor and learner were separated by location but not by time. With Internet-based distance education, a high degree of interactivity is encouraged among geographically separated learners. Learning is independent of time or place, using computer and communications technologies to work with remote learning resources, mentors, experts, and other learners, but without the requirement to be online at the same time (Staley & MacKenzie, 2000). Asynchronous group and individual messaging, as well as access to course materials and interactive realtime events, form the foundation for what constitutes online distance education today (O'Brien, 2001).

The Kirkpatrick (1994) four-level model of training evaluation shaped the research questions used for this multifaceted evaluation. The Kirkpatrick model consists of these four components: (1) Reaction—Were learners satisfied with the online courses? 2) Learning—Did learners' skills or knowledge increase or improve? (3) Behavior—Did learners apply skills and knowledge from online courses? (4) Results—Did online courses achieve desired outcomes of the school and do so cost-effectively?

This study examined K-12 online learning by asking and attempting to answer the following evaluation questions:

- Who enrolls in K-12 online learning courses, in what courses do they enroll, and why do they enroll for these courses?
- Is the achievement of students in teacher-directed online courses equal to or greater than their achievement in teacher-directed courses in high school classrooms?
- How are K-12 online courses implemented by students and teachers?
- What are the advantages and disadvantages of K-12 online learning?

Background and Review of the Virtual High School Movement

Distance education has been an educational option for students in the form of correspondence courses or interactive television since before the emergence of the Internet. The growth and popularity of the Internet and World Wide Web, however, has accelerated interest in distance learning (Roblyer, 1999). In the last ten years since the World Wide Web was founded, most educational institutions, libraries, government agencies, non-profit organizations, for-profit business enterprises, and many individuals have connected to the Internet. This proliferation of online communications has profoundly affected the growth of online training and education opportunities. Because online learning is a relatively recent development of distance education, much of the theory and research related to distance education in the form of online learning is still in the early stages of development.

Although Internet-based or online distance education has been more prevalent in post-secondary institutions, virtual high schools are starting to flourish on the Internet as an alternative delivery system for K-12 education (Berman & Tinker, 1997; Roblyer, 1999). Typically the rationale and justification for online learning in K-12 schools is different from that of higher education. Often online learning in K-12 schools focuses on expanding course offerings for students who are at risk, home-schooled, rural, or have disabilities, or provides access to educational resources and expertise unavailable in remote or rural areas.

Although virtual high schools have been in operation less than a decade, students continue to flock in record numbers to the "schools that technology built." Clark (2001) projected that as many as 50,000 students could be taking virtual courses by the end of 2002. He also found that the virtual high school trend was expanding rapidly throughout the U.S. educational system. Almost half of the 44 virtual high schools in his survey sample had been established in 2000-2001, a remarkable expansion of a phenomenon that only began around 1995.

Despite their ubiquity, virtual high schools defy easy definition as either an educational or technological phenomenon. The only common feature across all virtual high schools is a reliance on the Internet for at least some student-instruction interaction. Schools themselves vary greatly in administration, governing agencies, curriculum focus, instructional methods, and technologies. A brief description of these differing approaches will help supply a context for this study.

Organization and Governance of Virtual High Schools

Roblyer (2003) found that virtual schools have two kinds of administrative configurations: school coordination and school replacement. The most common arrangement is school coordination, which calls for individual schools to be part of a central organization such as the VHS, Inc. (formerly the Concord Consortium, see Zucker and Kozma, 2003). In this model, the consortium central office trains VHS instructors, organizes advertising and registration, and maintains a course Web site. However, course credit comes from the individual schools, each of which is housed in a physical school building and employs fulltime teachers who offer regular, face-to-face courses in addition to virtual ones.

The school replacement model also has centralized administration, but all courses are offered by an organization that is purely virtual, like the Florida Virtual High School (Bigbie & McCarroll, 2000). The organization obtains credit-granting authority from the county, state or country and offers all courses through its fulltime staff of teachers or facilitators. The vision for this latter model is that, eventually, we will be able to cut down on the number of physical school buildings, since students will not need to attend face-to-face classes at all.

Governance of virtual schools is more complex. Clark (2001) lists examples for seven different kinds of governing agencies that sponsor virtual schools or virtual courses for schools:

- **State-run** – Sanctioned by the state government and represented as the virtual school for the state (e.g., Florida Virtual School)
- **College/university** – Schools associated with higher education institutions and serve primarily to offer advanced courses to upper-division high school students through dual enrollment (e.g., University of Nebraska-Lincoln Independent High School)
- **Consortium/regional** – Multi-state, national, or international schools that act as brokers to provide courses for members through a variety of sources (VHS, Inc.)
- **Local education agencies** – Operate like consortia to broker courses for members but serve a more limited population, usually only the local state or area within a state (HISD Virtual School)
- **Charter schools** – Like other schools chartered by a state but run as online entities (e.g., Bashor-Linwood Virtual Charter School)
- **Private virtual schools** – Privately-funded schools established primarily to provide supplemental instruction and materials for home schooled students (e.g., WISE Internet High School)
- **For-profit providers** – Run by companies such as Apex Learning and Class.com, primarily to provide delivery platforms for many virtual schools and starter courses for new virtual high schools

Instructional Models and Technologies Used by Virtual High Schools

Although instructional methods used in virtual courses vary greatly, Roblyer (2003) finds that instructional approaches used by virtual schools can be classified according to three broad models: asynchronous interactive, online cyberbooks, and "live" or synchronous learning. Clearly, the asynchronous interactive model is the most common one. Courses use a system such as WebCT or BlackBoard to allow postings of materials by both students and instructors and discussions among them. These postings are supplemented by e-mail and chat exchanges, and most interaction takes place over time, rather than as live activities. Cyberbooks, or self-led tutorials consisting of posted materials with embedded web link sources, are much less common. In these, students go through instruction at their own pace, then take tests covering the concepts. Live interactive courses offered via some form of videoconferencing are also less common.

Virtual high schools also vary according to the primary type of delivery system they use. While all schools use the Internet for at least some purposes, some employ videoconferencing or broadcast technologies to support instruction and/or interaction. Loupe (2001) finds that some virtual schools still have a face-to-face component, which some schools seem to feel cuts down on drop-out rate and enhances student enjoyment of virtual courses.

Methodology

Overview and Description of the Virtual High School

The subject of this evaluation was a virtual high school operated by an education services cooperative in a midwestern state in the United States. This organizational model of the virtual high school was school coordination. The virtual high school targeted school districts with a need for building educational capacity through the delivery of online courses that expand curriculum offerings, eliminate scheduling conflicts, address teacher shortages, and cost-effectively enable a school district to extend learning opportunities. Courses offered by the virtual high school included some it developed as well as many courses *brokered* from content providers. Each course provided by the virtual high school was extensively evaluated utilizing a validation rubric so that it can be integrated into a district curriculum.

Effective partnerships with online content providers were an important factor for the success of this virtual high school. For this study, 86% of student enrollments were in courses provided by Class.com; 5% were in courses provided by Apex Learning; 3% were in courses provided by Cyberschool; and 6% were in courses developed by the virtual high school.

Role of Teachers in Online Courses

All the online courses had a teacher designated for the course and were more or less teacher-directed. Teachers were paid based on the number of students that enrolled in their course. Teachers received a contract on a new enrollee once the student had cleared the 14-day period in which it was possible to drop a course and receive a refund. There was no imposed minimum or limit on the number of students a teacher had in a course. For the period of time covered by this evaluation, the average course size was 122 students, although actual course sizes varied considerably from as few as 5 students to over 250 students.

Teachers were expected to contact new enrollees by email as they enrolled in a course to provide some information about the teacher along with the expectations the teacher had for student progress in the course. Teachers were required to publish office hours that they are available to provide immediate assistance to students and were required to be available for a minimum of one hour, one day per week. Teachers were expected to check email accounts and discussion boards daily and to update grades daily. Grades were assigned using a percent basis, permitting the local school district to grant the credit and assign the letter grade.

Role of Students in Online Courses

The virtual high school provided online courses that allow students to take online courses for high school credit for transfer to a local school district. Students enrolled in online courses were expected to abide by the virtual high school's Acceptable Use Policy (AUP). Upon enrollment students were required to sign a Student Contract, indicating they would comply with the AUP. The Student Contract also established procedures for removal from a course in the case of non-compliance with the AUP.
Students had six months to complete a .5 unit of credit course and twelve months to complete a 1.0 unit of credit course. Completion time for a course was calculated from the date the student enrolled in an online course. A student could complete a course, however, in as few as 30 days. In its analysis of course completions, this evaluation did not distinguish between half credit and full credit courses nor did it attempt to analyze course completion and achievement by curricular subject or theme.

Study Sample and Analysis Methods

The primary unit of evaluation and analyses for this study was student enrollments (N=2601) in online high school courses. The count of course enrollments was not unduplicated, and so most analyses computed cases based on students in each online course in which they enrolled and not as individual students. From the total, student enrollment frequency counts were used to derive descriptive statistics for the study sample. To examine student performance and achievement in online high school courses, a cohort of course completers was extracted from the enrollment database. The course completers cohort (N=702) consisted of all student enrollments who completed a course prior to the commencement of the data collection for this study.

To compare student achievement in school with achievement in online courses, student cases were randomly selected from the course completers cohort. Letters and forms were mailed to the parents/guardians of the randomly selected completers. Using this random selection process, permission forms from parents and grade records from schools were obtained for only 27 students (including 1 home school student). Because the response to this solicitation did not provide a large enough sample for analysis, school districts with substantial enrollments in the virtual high school were invited to provide transcript and standardized test score information for students enrolled in online courses. Transcript and test score information was obtained for an additional 40 students to bring the total to 67 student cases with both online course grades and school achievement records.

To evaluate student perceptions of course quality and delivery, students completing an online course were requested to complete an online course evaluation. Valid and usable forms from 307 student evaluations were examined. Student evaluation forms, like enrollment records, were not unduplicated and so the analyses computed cases based on valid, completed forms and not as individual students.

Data Analysis

Original data sources in the form of the student enrollment database and school transcripts were examined for this evaluation. Of the student enrollments examined for this study, 46.3% were female, 52.2% were male, and 1.5% were not disclosed. The distribution of enrollments by content area was as follows: science (20.6%), foreign languages-mostly Spanish (16.4%), mathemathics (16.1%), social studies/social sciences (14.5%), language arts (13.2%), health (9.2%), skills development—for example, keyboarding (4.2%), business, (3.2%) art (1.3%) and technology (1.3%).

High School Student Achievement in K-12 Online Courses

To examine the achievement of students in K-12 online courses, all students who completed an online course during the evaluation period were extracted from the enrollment database. The number of days to complete the course, the number of days in which a course was completed prior to maximum completion, and the mean and median days for completion and for course grades were computed for all course completers (see Tables 1 and 2). Completion rates by content provider and online course grades by subject for all completers are listed in Tables 3 and 4.

Table 1: *Course Completion Rate in Days and Final Course Grade (6 month courses, N=530)*

Time/Achievement Variable	***Minimum***	***Maximum***	***Mean***	***Median***
Days to complete	15	440	155.1	140.0
Days before maximum completion	0*	310	29.6	45.5
Grade in online course	23	100	85.3	88.0

**Some students went beyond the maximum completion date to complete the course.*

Table 2: *Course Completion in Days and Final Course Grade (12 month courses, N=172)*

Time/Achievement Variable	***Minimum***	***Maximum***	***Mean***	***Median***
Days to complete	27	466	162.9	140.0
Days before maximum completion	0*	338	192.3	225.0
Grade in online course	57	99	85.5	88.0

**Some students went beyond the maximum completion date to complete the course.*

Table 3: *Completion Ratios by Course Provider (N=702)*

	Enrollment		**Completions**		
Provider Name	**Frequency**	**Percent**	**Frequency**	**Percent**	**Completion Ratio**
Apex Learning	116	4.4	0	0	0%
Class.com	2260	86.9	658	93.7	29%
CyberSchool	75	2.9	3	.4	4%
eBush Learning	150	5.8	41	5.8	27%
Total	**2601**	**100.0**	**702**	**100.0**	**27%**

Table 4: *Online Course Grades for All Course Completers*

Subject/Category	***N***	***Grade Mean***	***Grade Std. Deviation***	***Scale Mean***	***Scale Std. Deviation***
Mathematics	91	82.76	11.99	2.80	1,22
Science/Health	285	84.25	10.92	3.00	1.94
Language Arts/Foreign Language	122	86.14	9.32	3.27	.79
Social Studies/Social Science	105	87.34	9.94	3.29	.94
Other	99	89.33	7.30	3.51	.71
All Subjects	702	85.56	10.40	3.14	.99

The correlation between student grades in the online courses and measures of achievement in school—Cumulative GPA and ACT Composite—were examined in this evaluation (see Tables 5 and 6). Because the sample size was relatively small (N=67), all cases in the sample were used for analysis and no tests of normality were conducted. To conduct this analysis, it was necessary to convert both course grades and transcript grades for each of the students in the sample to a similar scale. Assigning a letter grade (or transcript grade) to an online course numerical grade is the responsibility of the participating school district where the student resides. In most cases this analysis used the actual scale used by the district assigning the letter grade. When the actual grading scale was unknown for a student case, a scale used by most school districts in the sample was used to assign a letter grade: A=90-100, B=80-89, C=70-79, D=60-69, F=less than 60. Next, the course letter grades and transcript letter grades were assigned a numerical value using the scale: A=4, B=3, C=2, D=1, F=0. The Spearman rho was used to compute the correlation between the online course grade and other school measures of achievement because it is appropriate for data that do not satisfy the normality assumption. The Spearman rho yielded significant correlations ($p < .000$) between the online course grade and the cumulative GPA and ACT Composite score.

Table 5: *Student Achievement from Student Transcripts and Online Course*

Measures of Achievement	***N***	***Mean***	***Std. Deviation***
Online Course Grade	67	82.567	9.545
Cumulative GPA	67	3.301	.723
ACT Composite	39	23.770	3.700

Table 6: *Correlation between Online Course Grade and School Measures of Achievement*

School Measures of Achievement	***N***	***Sig. (2-tailed)***	***Correlation Coefficient***
Cumulative GPA	67	.000	.524
ACT Composite	39	.000	.548

High School Student Attitudes about Online Learning

One component of the evaluation process consisted of collecting feedback from students concerning their attitudes about taking online courses. An online student evaluation form was developed to collect these data. The online form permitted the virtual high school students to complete an online evaluation by linking to the form once they completed their course. Of the students who completed an online evaluation form, 60% were female and 37% were male (3% not disclosed). The grade level of students taking an online course was widely distributed across high school grade levels: 31% were seniors, 28% were juniors, 25% were sophomores, 11% were freshman, and 5% were 7th or 8th graders. Most students (83%) completing an online evaluation form reported they had never taken on online course before. Most students who evaluated an online course accessed the course from school (72%) or from home (25%) and 87% owned a home computer. Most students self-reported that they made above average grades in school with 46% of the students reporting mostly A's in school, 29% mostly B's, 17% mostly C's, and less than 8% with mostly D's or F's or not disclosed.

There were 20 positive statements about the quality or usability of some feature of the online course presented in the online evaluation form. Students were asked to react to the statements by rating each statement along a Likert scale as Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree.

For analysis purposes a value was assigned to each possible response along an ordinal scale from –2 to + 2 where Strongly Disagree was –2, –1 was Disagree, Neutral was 0, Agree was +1, and Strongly Agree was +2. Table 7 provides the mean student rating on this scale for the top 6 statements (Mean >.50) on the evaluation form by degree of agreement with the statement. The table lists the statements by rank from those statements receiving the highest mean rating of agreement to those receiving the lowest mean rating of agreement.

Table 7: *Mean Rankings of Evaluation Statements (N=287)**

RANK (high to low)	**MEAN**	**STUDENT EVALUATION PROBE**
1	.79	16. The instructor seemed to have a thorough knowledge of the subject.
3	.65	14. The instructor provided helpful feedback and assistance throughout the course.
2	.61	15. The instructor encouraged me to communicate with her throughout the course.
4	.61	13. The instructor was sensitive to student difficulties with the course work.
5	.59	9. The work and study requirements and grading system were made clear and were fair.
6	.53	10. Course assignments were challenging, worthwhile, and helped me learn about the subject.

Scale: -2=Strongly Disagree, -1=Disagree, 0=Neutral, +1=Agree, +2=Strongly Agree

The student responses for each of the 20 items were summed to create a total scale score indicating overall student satisfaction with the online course. The mean for the total scale was 7.87. A stepwise regression model was formulated to determine those statements that best predicted the overall student satisfaction with the online course. Four factors predicted over 90% of the variance in student responses on the scale (see Table 8).

Table 8: *Best Predictors of Overall Student Satisfaction with an Online Course (N=287)**

PREDICTOR	**R**	**R Square**	**R Square Change**
20. I would recommend this course to others.	.841	.707	.707
14. The instructor provided helpful feedback and assistance throughout the course.	.919	.844	.137
10. Course assignments were challenging, worthwhile, and helped me learn about the subject.	.943	.888	.044
4. I was satisfied with the amount of interaction I had with the Web course teacher.	.956	.915	.026

Discussion

Overall, online delivery of instruction fulfills a very real and practical need in the high school curriculum, especially for students in small or rural high schools. While some limitations on course offerings exist in all high schools, in those schools restricted by access to resources and/or expertise due to geography or socio-economic factors, curriculum-based, online courses for high school students may provide a realistic and affordable solution for these limitations in the curriculum of many secondary schools. Although students often reported that they enrolled in online courses due to scheduling issues—the course was not offered at their school, there was a scheduling conflict, or they were making up credit—students also were likely to enroll in online courses for reasons of enrichment or to expand their learning.

Having quantified and qualified many sources of data for this evaluation, it is now appropriate to ask: What did this evaluation reveal about this virtual high school's programs that we did not know before? The following findings are based on general interpretations of the data collection and analysis performed for this evaluation. These findings, hopefully, provide some insight into the design and delivery of online learning in a virtual high school environment:

- Online learning, even teacher-directed online learning, operates under the supposition that students are (or want to be) responsible for their own learning. High school students in general, however, are not equipped to be responsible for their own learning (and therefore may not want to be responsible).

Possibly, this incapacity of high school students is the effect of many years of learning experiences in a public education system that tells students what, when, where, and how much (or how little) to learn.

- The primary problems and issues related to online learning are technical difficulties related to access or installation. Some technical problems may occur at the local level due to old hardware or slow Internet connections, while other technical problems may occur at the delivery level (e.g., students not understanding how to properly install or use course, course developers not implementing universal design principles into course content).
- A large number of students who enroll in online courses ultimately drop out. A number of reasonable explanations for drops exist that are not necessarily related to course quality, design, or delivery problems or issues. In many studies, attrition rates are cited as a major difficulty of online education (see Roblyer & Elbaum, 2000). It is important to recognize that because online courses are anytime/anywhere, that characteristic itself may give support to the likelihood of many drops. Therefore, the virtual high school should be equipped with a student feedback process for debriefing students who drop courses (as well as those who complete courses).
- One feature of online learning that proponents usually tout as its main strength is asynchronous communications. High school students, however, often perceived this feature as a disadvantage of online learning. High school students seemed to consider learning as consisting of a series of *teachable moments* that are created by something a teacher says or does in the classroom.
- Students enrolled in online courses for reasons of enrichment as much as matters of necessity. Although students often enrolled in online courses to solve scheduling conflicts or to make up credits, almost as often students enrolled in online courses because they were interested in the subject and the courses were not offered at their school.
- Although students often admitted to procrastinating or being undisciplined in their approach to completing online learning activities, when students did complete a course, they generally completed it in a reasonable amount of time.
- Students had favorable attitudes about online teachers. Teachers played an important role in online instruction. The synchronous and asynchronous interventions they provided were an important component of online teaching and learning.
- Based on results from a small sample of students, it appeared that student achievement (grades) in online courses was at least equal to (and slightly) higher than student achievement in high school classrooms.
- Online courses have the potential for enriching the curriculum of a school and increasing the learning opportunities for students.
- Online courses provide a standards-based curriculum resource for home-schooled students.

Recommendations for Improvement and Future Research

The following recommendations for continuous improvement of the virtual high school are proffered, based on the findings of this evaluation. (Note: Many of the following recommendations have been instituted by the virtual high school since this evaluation study was completed.)

- Administer a diagnostic assessment to students enrolling in an online course to help predict their success in online courses and provide an online orientation to reinforce the results of the diagnostic assessment.
- Collect additional data on the reasons why students drop online courses by developing a system for conducting exit interviews or obtaining feedback from students who drop online courses.
- Focus more evaluation efforts and analytical methodologies on individual course design and course design by provider. A sample of high and low enrollment and completion courses should be selected for instructional analysis to identify features of course design that may promote or hinder online learning. These analyses should facilitate comparisons between the design and implementation features of online courses and educational best practices for teaching and learning with technology as well as ensure that online courses incorporate principles of universal design for education.
- Place higher expectations for course completion on students by decreasing the amount of time a student has for completing a course. The features of online learning should be fully exploited to realize optimal efficiency in online course-taking and completion. Higher expectations on students may also serve to increase the discipline and responsibility of the individual student for learning in an online instructional environment.

- Increase the amount of synchronous and asynchronous communication between online instructors and students by increasing online "office hours" and imposing limits on course size per instructor. Students relate learning to the presence of a teacher and at times may not proceed with online learning activities until a teacher resolves the issue. Additionally, a process for logging and analyzing synchronous and asynchronous instructor/student interactions should be implemented.
- Decrease course size to permit teachers to provide more online time with each student.
- Implement agreements with participating high schools to provide student transcripts to use for comparisons between online course achievement and achievement in school. Participating schools should provide profiles of the grading scale and the units of credit required for high school graduation .
- Review the curriculum of courses from providers with low student completion rates. Minimum standards for student completion rates should be maintained for all course providers. Providers should be informed of student completion rate standards and when rates fall below the minimum, the provider curricula should be subject to review to determine the reason for low completion rates.

References

Berman, S., & Tinker, R. (1997). The world's the limit in the virtual high school. *Educational Leadership, 55*(3), 52-54.

Bigbie, C., & McCarroll, W. (2000). *The Florida high school evaluation 1999-2000 report.* Tallahassee: Florida State University.

Clark, R.E. (2001). New directions: Evaluating distance education technologies. In R.E. Clark (Ed.), *Learning from Media.* Greenwich, CT: Information Age Publishing.

Clark, T. (2001, October.) *Virtual Schools: Trends and Issues*, report commissioned by the Distance Learning Resource Network, a WestEd Project; co-sponsored by the Center for the Application of Information Technologies at Western Illinois University. [Available online: http://www.dlrn.org/trends.html]

Kirkpatrick, D.L. (1994). *Evaluating training programs*. San Francisco: Berrett-Koehler.

Loupe, D (2001) Virtual schooling: A new dimension to learning brings new challenges for educators, *eSchool News, 4*(6), 41-47

O'Brien, F. (2001). *Learning style theory: Its implications for the design and delivery of on-line distance education courses.* Trinity College, University of Dublin, Dublin, Ireland.

Roblyer, M.D. (1999). Is choice important in distance learning? A study of student motives for taking Internet-based courses at the high school and community college levels. *Journal of Research on Computing in Education, 32*(1).

Roblyer, M. D. (2003). Virtual high schools in the United States: Current views, future visions. In Jo Bradley (ed.), *The open classroom: Distance learning in and out of schools.* London: Kogan Page.

Robleyr, M.D., & Elbaum, B. (2000). Virtual learning? Research on virtual high schools. *Learning & Leading with Technology, 27*(4), 58-61.

Staley, A., & MacKenzie, N. (2000). Enabling curriculum re-design through asynchronous learning networks. *Journal of Asynchronous Learning Networks, 4*(1).

Zucker, A., & Kozma, R. (2003). *The virtual high school.* New York: Teachers College Press.

Students' Perceived Ease of WebCT's Use: An Exogenous or Endogenous Variable?

Cheng-Chang (Sam) Pan
Stephen Sivo
Glenda Gunter
James Brophy
Edward Hampton
Richard Cornell
University of Central Florida

Abstract

Five factors affecting student use of a course management system in two Web-enhanced hybrid undergraduate courses are investigated using the technology acceptance model (TAM). This research represents a causal relationship existing between students' attitude toward WebCT and their actual use of the system. Students' perception of the WebCT use, computer self-efficacy, and subjective norms are also taken into account. Multigroup structural modeling procedure, specifically PROC CALIS, is used to extract those factors from student use of WebCT and to determine their inter-relatedness among one another. Results show that extended adaptations of the Technology Acceptance Model are not as suitable for Engineering students as they are for Psychology students. Of the two competing models in the psychology class, Perceived Ease of Use is deemed an exogenous variable. A multi-sample analysis suggests that covariance structure differences between psychology and engineering students were found obvious over Computer Self-Efficacy and Subjective Norms variables. Lessons and experience from a southeastern metropolitan university in the U.S. are addressed.

Introduction and Background of the Study

To meet diverse needs of the student body, Web-enhanced classes using WebCT are currently offered at the University of Central Florida (UCF). In the present study, WebCT is conceptualized as an information system project and it is also considered a course management system. This study was concerning the inter-relationship among students' perception of WebCT design features, their attitude toward WebCT, and their WebCT use.

Fishbein and Ajzen's (1975) theory of reasoned action (TRA) specified a causal relationship between individual behavioral intention and actual behavior. With TRA, one can differentiate an individual's actual behavior from his or her behavioral intention. Behavioral intention is a latent factor that is measured by two other latent factors: attitude toward behavior and subjective norm.

Rooted in TRA, the Technology Acceptance Model (TAM) by Davis (1989) identifies two distinct constructs, perceived usefulness and perceived ease of use. Those two constructs directly affect a person's attitude toward the target system use and indirectly affect actual system use (Davis, 1993). Davis (1993) defined perceived usefulness as "the degree to which an individual believes that using a particular system would enhance his or her job performance" and perceived ease of use as "the degree to which an individual believes that using a particular system would be free of physical and mental effort" (p. 477). Furthermore, attitude toward use of a system is defined as "the degree to which an individual evaluates and associates the target system with his or her job" (p. 476). Actual system use is a behavioral response, measured by the individual's actions in real life. David (1993) states that "frequency of use and amount of time spent using a target system are typical of the usage metrics" (p. 480).

The TAM is used by Management Information Systems (MIS) practitioners to predict the success or a failure of an information systems project. The TAM is based on the following assumptions:

1. When end users perceive the target system as one that is easy to use and nearly free of mental effort, then they may have a favorable attitude toward using the system. Nevertheless, Sanders and McCormick (1993) argued that an individual must use some of or all of one's mental resources in order to perform a task.
2. When end users perceive the system as one that is helpful to their job, then they may have a positive

attitude toward the system used.

3. When users have a favorable attitude toward the target system, they may use the system frequently and intensely, which means that the system developed is successful.

Above all, the TAM was adapted to predict the acceptance or rejection of WebCT by the participating classes when the courses go fully Web-based.

Purpose and Relevance of the Study

The first phase of this correlational study was to investigate the causal relationship existing among student perception of WebCT, student attitude toward the use of WebCT, their actual system use and two other external variables: Subjective Norms and Computer Self-Efficacy. In doing so, the authors successfully replicated the Technology Acceptance Model (TAM) and extended TAM in a higher education setting by verifying a belief-attitude-behavior relationship in the context of WebCT adaption (Pan, 2003).

Because two data sets collected at the two time occasions were analyzed separately in the first phase of analysis, in the second phase of the analysis the authors succeeded to achieve the plausibility of a hypothetical TAM by using five factors from Time 1 data: Subjective Norms, Computer Self-Efficacy, Perceived Ease of Use, Perceived Usefulness, Student Attitude toward WebCT's Use and three outcome variables from Time 2 data: Frequency, Intensity, and Grade or student end-of-course achievement (Pan, Sivo, Gunter, & Cornell, under review).

The primary purpose of the present study (Phase Three) is two-fold. First, they attempted to verify the role of students' perceived ease of WebCT's use in the presence of two external variables: Subjective Norms and Computer Self-Efficacy. Second, they strived to determine differences of the factor covariance structures between the two student groups by conducting a multi-sample analysis using structural equation modeling.

The significance of this study may provide administrators with insights into student perception about the system employed and their individual traits, which may mediate the acceptance of such technology. Combining these components with the findings from previous WebCT-related studies (e.g., Dziuban & Moskal, 2001; Moskal & Dziuban, 2001) can assist decision makers in analyzing the influence of WebCT on varied UCF campuses, including physical and virtual ones.

Review of Literature

The Technology Acceptance Model

The Technology Acceptance Model (TAM) originated from the psychological environment and expanded into the business settings. Adapted from the Theory of Reasoned Action (TRA), the Technology Acceptance Model (TAM) by Davis (1989) identified two distinct constructs, perceived usefulness and perceived ease of use, which directly affect the attitude toward target system use and indirectly affect actual system use (Davis, 1993). Each of the factors is defined as follows:

- Perceived ease of use: the degree to which the individual believes that using the target system would be free of mental and physical efforts (Davis, 1993, p. 477).
- Perceived usefulness: the degree to which an individual believes that use of the target system could enhance the job performance (Davis, 1993, p. 477).
- Attitude toward use of target system: the degree to which an individual evaluates and associates the target system with his or her job (Davis, 1993, p. 476).
- Actual system use: a behavioral response, measured by the individual's action in reality (Davis, 1989).

The causality of the four components of the Technology Acceptance Model addressed previously can be explained theoretically. Reversely speaking, management Information Systems (MIS) research bases the success of actual system use on the frequency and intensity of the target system use (Davis, 1993). Attitude measures the tendency toward actual system use (e.g., Davis, 1985; Harris, 1999; Lu, Yu, & Lu, 2001). According to Davis (1989), when the causal relationship between attitude and usage is established, then antecedents or determinants of end user attitude toward the target system are not as difficult to investigate. The antecedents mentioned referred to end-user perception about the easiness and usefulness of the IT system.

From a system design features' viewpoint (Davis, 1985), the TAM identified two vital determinants of end users' attitude toward the technology: perceived ease of use and perceived usefulness. The causal relationship of perceived ease of use to perceived usefulness is corroborated by Hubona and

Blanton (1996). Hubona and Blanton measured the predictive capabilities of perceived ease of use and perceived usefulness to three variables: task accuracy, task latency (i.e., response time), and user confidence in decision quality; their findings suggested that users' perceived ease of use affects the three outcome variables much more significantly than users' perceived usefulness. This is supported by Igbaria, Zinatelli, Cragg, and Cavaye (1997), who demonstrated that administrative/management support coupled with external expert support (e.g., vendors) can influence perceived ease of use and perceived usefulness, which, in turn, contributes to system use.

The effect and power of the TAM in predicting user acceptance have been validated through almost two decades. A meta-analysis of the Technology Acceptance Model by Legris, Ingham, and Collerette (2003) reported three major findings:

1. Varied settings were employed in terms of the target system and the TAM has been cross-examined and compared to other theories.
2. Researchers adapted the TAM to a varying degree and they focused on different relationship among variables.
3. External variables used varied from one study to another.

Legris, et al (2003) cited the target systems used in the TAM studies could be grouped into three categories: office automation tool, software development tool, and business application tool. Their review indicated that those TAM researchers either compared the TAM to the theory of reasoned action (TRA) or the theory of planned behavior (TPB) and some researchers added subjective norms construct to their expanded TAM-based model, a variable examined in the TRA. With respect to the versions of the TAM, the twenty-two TAM-based studies reviewed by Legris et al. delineated a high proportion of congruent findings on each of the ten possible linear relationships among variables. Regarding the external variables measured in the TAM studies, Legris, et al. continued that variables such as computer self-efficacy and subjective norms were added to the hypothetic model together with others, which allowed for a fuller picture of the model. Though the antecedents accounted for trivial increases on the explained variances of the outcome variable, the presence of those external variables afforded strategic plans for more likelihood of user acceptance. Furthermore, attitudes toward the system use and behavioral intention to the system use appeared to be either used interchangeably or overlooked in the literature, while some studies included both variables. Actual system use, the outcome variable in the TAM, was either measured, using frequency and duration of system use, or disregarded by the researchers.

External Variables

Originally from Bandura's (1977) self-efficacy theory, computer self-efficacy becomes a pivotal issue in technology acceptance. Venkatesh and Davis (1994) first coined the term computer self-efficacy, which is defined as "[the] self-efficacy…in the specific context of user acceptance of computer technology" (p. 214). Venekatesh and Davis reported that users' perceived ease of use is strongly regressed on computer self-efficacy in the early stage of technology acceptance. To their convenience, the authors used computer self-efficacy in the present study to denote self-efficacy for online learning systems skills in Web-enhanced courses.

Subjective norms include users' perception of the external forces and their motivation to comply with the forces (Robinson, 2001). Wolski and Jackson (1999) endorsed this proposition from the perspective of university faculty in the context of faculty development.

Research Methodology

This study is a research investigation using structural equation modeling. Derived from Davis' (1985) Technology Acceptance Model (TAM), two hypothetic models were designed to compete with each other in order to verify the role of students' perceived ease of WebCT's use in the presence of two external variables: Computer Self-Efficacy and Subjective Norms and to measure the factor covariance structure differences between the two classes of student participants: the psychology class and the engineering class. The purpose of this study was to answer the following questions:

1. Does student perceived ease of WebCT's use variable remain an exogenous variable in the presence of the two external variables?
2. To what extent does the psychology class differ from the engineering class with respect of factor covariance structures involved in the study?

Design and Sample of the Study

This is a correlational research study of students' use of WebCT in two WebCT-enhanced undergraduate courses in the University of Central Florida (UCF) in Orlando, Florida. This structural equation modeling study with quantitative measurements concentrated on the Web-enhanced hybrid courses, particularly the two large-sized undergraduate courses: PSY2012a General Psychology course and EGN1007a Engineering Concepts and Methods. In the psychology class, 230 out of 239 participants were randomly selected. In the engineering class, all of the 230 participants were included in the analysis.

In this paper, the causality issue in the belief-attitude-behavior relationship was scrutinized from the students' perspective of WebCT use in the WebCT-enhanced hybrid courses across one semester with an emphasis on students' perceived ease of WebCT's use. Given this context and based on the previous findings, causal pathways among Students' Perceived Ease of Use, Perceived Usefulness, Attitudes toward Using WebCT, their personal Subjective Norms, Self-Efficacy with regard to WebCT, and Actual Use of WebCT were re-explored.

Data Collection and Analysis

Endorsed by the University of Central Florida Institutional Review Board, an online questionnaire with seven varied scales was administered to students in the two courses on two time occasions in the Spring semester of 2003. The instruments included (1) a Usability Instrument (including perceived ease of use and perceived usefulness scales); (2) an Attitude Instrument; (3) a Computer Self-Efficacy Instrument; (4) a Subjective Norms Instrument; (5) a WebCT Use Instrument; (6) a Student Demographic Instrument. For the scope of this study, all the variables from the first five instruments were analyzed, which yielded a total of 51 variables (including student achievement variable) for each class. To conduct a categorical analysis to determine the factor covariance structure differences between the two groups as suggested by Marcoulides and Hershberger (1997, p. 252), 102 variables were taken into account in this study. In acknowledging the linear dependency between variables found at the bottom level (Pan, Sivo, & Brophy, in press; Sivo, Pan, & Brophy, in press), the authors remained to conduct this study on a factor level.

Data analysis of the present study was composed of two stages: testing the two competing models on both classes separately and examining the factor covariance structure differences between the two classes. After sampling the same number of participants in the two classes, a SEM procedure, PROC CALIS, was used to model all the variables and error terms at one time on a scale level. Then, taking Marcoulides and Hershberger's (1997) suggestion, the authors sought to "fool" the SAS program to undergo a factorial analysis of the two groups, using PROC CALIS, as opposed to EQS or LISREL. PROC PRINT was used to generate covariance matrices prior to the factorial analysis. The debate over use of PROC CALIS in the categorical analysis is beyond the scope of this SEM study.

The following fit indices were examined: Comparative Fit Index (CFI) and the Standardized Root Mean Square Residual Estimate (SRMR). These indices were chosen because of their relative merits. The CFI is an Incremental Fit Index that indicates how much the fit of a model improves upon the nested null model. This index is more sensitive to misspecification between latent and manifest variables relationship misspecifications. The SRMR is more sensitive to latent-latent variable relationship misspecifications.

An assessment of adequate fit in structural equation modeling is not without standard cutoff criteria. In part, the cutoff criteria chosen are the result of Hu and Bentler's (1999) monte carlo simulation findings. The CFI is expected to exceed .95 if the model is to be deemed as fitting well. The SRMR is expected to attain values no higher than .05.

Results

The descriptive statistics calculated for the two groups presented in Table 1 reveal that both groups were overall similar in terms of the variable means and standard deviations.

Table 1. *Descriptive Statistics for Student Participants*

Variable	Mean	Standard Deviation
	Psychology Class (n = 230)	
Perceived Usefulness of WebCT (PU_T1)	30.857	6.731
Perceived Ease of WebCT Use (PEU_T1)	33.522	6.801
Attitude regarding WebCT Use (AT_T1)	28.822	5.136

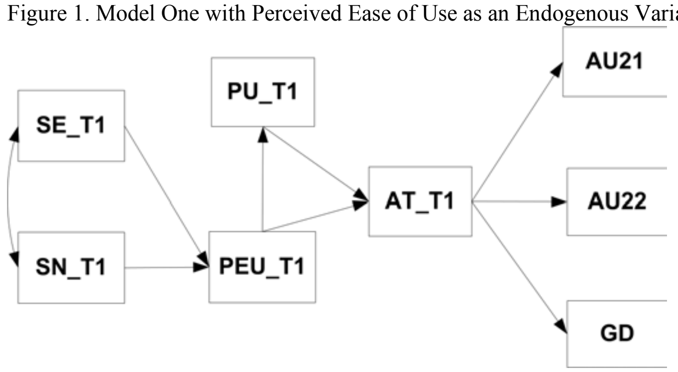
Self-Efficacy (SE_T1)	169.609	20.252
Subjective Norms (SN_T1)	21.604	3.553
Frequency of WebCT Use (AU21)	4.522	0.704
Intensity of WebCT Use (AU22)	2.061	0.889
	Engineering Class (n = 230)	
Perceived Usefulness of WebCT (PU_T1)	29.613	6.594
Perceived Ease of WebCT Use (PEU_T1)	33.522	7.080
Attitude regarding WebCT Use (AT_T1)	27.504	5.102
Self-Efficacy (SE_T1)	172.730	23.149
Subjective Norms (SN_T1)	20.961	3.387
Frequency of WebCT Use (AU21)	3.330	1.273
Intensity of WebCT Use (AU22)	2.135	0.982

Although the Engineering students scored a little higher than the Psychology students with respect to Self Efficacy in terms of computer use, the difference is not pronounced enough to warrant attention, especially given the variability in the scores, as suggested by the standard deviations.

Research Question One: Does student perceived ease of WebCT use remain an exogenous variable in the presence of the two external variables?

To answer this question, two structural models, serving as rival hypotheses were fitted to the covariance data for psychology and engineering students separately (see Figure 1 and Figure 2).

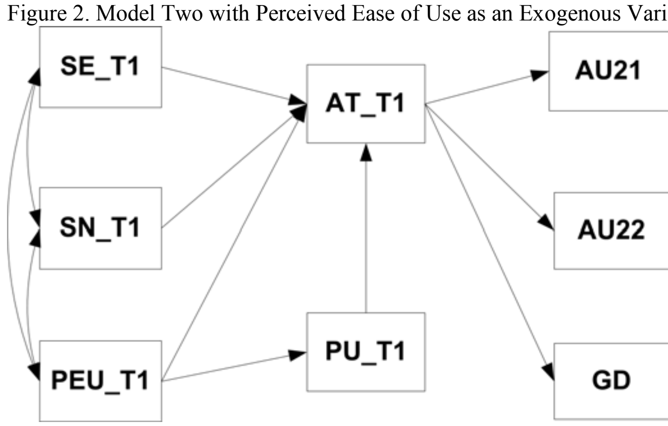
Figure 1. Model One with Perceived Ease of Use as an Endogenous Variable



Note. Note. SE_T1= Computer Self-Efficacy at Time 1; SN_T1= Subjective Norms at Time 1; PEU_T1= Perceived Ease of Use at Time 1; PU_T1= Perceived Usefulness at Time 1; AT_T1= Attitude toward WebCT at Time 1; AU21= Frequency at Time 2; AU22= Intensity at Time 2; GD= Grades.

The two extended versions of the Technology Acceptance Model were pitted against one another: Model One specifying perceived ease of WebCT use as endogenous; Model Two, as exogenous.

Figure 2. Model Two with Perceived Ease of Use as an Exogenous Variable



Note. SE_T1= Computer Self-Efficacy at Time 1; SN_T1= Subjective Norms at Time 1; PEU_T1= Perceived Ease of Use at Time 1; PU_T1= Perceived Usefulness at Time 1; AT_T1= Attitude toward WebCT at Time 1; AU21= Frequency at Time 2; AU22= Intensity at Time 2; GD= Grades.

The maximum likelihood procedure was able to successfully converge upon a proper solution for all models fitted to the psychology and engineering covariance data. A review of the fit statistics revealed that the rival models fitted the psychology student covariance data well, but did not do the same with the engineering student covariance data. These findings suggest that both models explain the pattern of responses collected from the psychology students well, but do not explain responses given by engineering students.

Table 2. *Fit results for Rival Models by Student Group* [a]

Fit Index	Model One: Perceived Ease of Use as Endogenous	Model Two: Perceived Ease of Use as Exogenous
	Psychology Class	
Bentler's Comparative Fit Index (CFI)	0.9675	0.9797
Standardized Root Mean Square Residual (SRMR)	0.0587	0.0490
Chi-Square	29.9268	23.8224
Chi-Square DF	19	17
Pr > Chi-Square	0.0527	0.1243
	Engineering Class	
Bentler's Comparative Fit Index (CFI)	0.8239	0.8709
Standardized Root Mean Square Residual (SRMR)	0.0969	0.8000
Chi-Square	86.2445	66.2865
Chi-Square DF	19	17
Pr > Chi-Square	<.0001	<.0001

[a] Covariance Structure Analysis: Maximum Likelihood Estimation

Because neither Model One nor Two fitted the Engineering student data, further attention will be confined to the Psychology student data. When comparing the fit of Model One and Two to the psychology student data, the difference in fit results is slight, although if Hu and Bentler's (1999) standards for these data were rigidly applied Model Two would be preferred as the SRMR for Model One exceeds the criterion of .05 (at .0587). Moreover, the chi-square statistic for Model One has an associated probability of .0587, although not statistically significant at the .05 level. Beyond these considerations, the coefficients of the paths should be evaluated, with particular attention given to Perceived Ease of Use.

Table 3 presents the coefficients associated with Model One's structural equations.

Table 3. *Manifest Variable Equations with Estimates for Model One fitted to Psychology Student Data* [a]

```
PAU21  = 0.00535*PAT_T1  + 1.0000 e1
Std Err   0.00905 GA1
t Value   0.5910

PAU22  = 0.0122*PAT_T1  + 1.0000 e2
Std Err   0.0114 GA2
t Value   1.0662

PGD    = 0.0172*PAT_T1  + 1.0000 e3
Std Err   0.00741 GA3
t Value   2.3246

PAT_T1 =  0.4326*PPU_T1  + 0.1200*PPEU_T1  + 1.0000 e4
Std Err   0.0457 GA4       0.0452 GA6
t Value   9.4750           2.6556

PPU_T1 =  0.5680*PPEU_T1  + 1.0000 e5
Std Err   0.0536 GA5
t Value   10.6048

PPEU_T1 =  0.1457*PSE_T1  + 0.2177*PSN_T1  + 1.0000 e6
Std Err    0.0210 GA7       0.1195 GA8
t Value    6.9451           1.8212
```

[a] PPU_T1=Perceived Usefulness; PPEU_T1=Perceived Ease of Use; PAT_T1=Attitude; PGD=Grades PSE_T1=Self-efficacy; PSN_T1=Subjective Norms; PAU21=Frequency of Use; PAU22=Intensity of Use

The results reveal that neither Frequency of Use nor Intensity of Use are predicted very well by Attitudes towards WebCT use, although student final grades were predicted to statistically significant degree, though a very small degree. Table 4 indicates that only 2.3% of the variation in student grades was explained by Model One.

Table 4. *Squared Multiple Correlations for Model One fitted to Psychology Student Data* [a]

Variable	Error Variance	Total Variance	R-Square
PAU21	0.49425	0.49500	0.00152
PAU22	0.78709	0.79100	0.00494
PGD	0.33118	0.33900	0.0231
PAT_T1	14.50291	26.37400	0.4501
PPU_T1	30.38769	45.31100	0.3294
PPEU_T1	35.19718	46.25100	0.2390

[a] PPU_T1=Perceived Usefulness; PPEU_T1=Perceived Ease of Use; PAT_T1=Attitude; PGD=Grades PSE_T1=Self-Efficacy; PSN_T1=Subjective Norms; PAU21=Frequency of Use; PAU22=Intensity of Use

On the other hand, the variation in scores for Perceived Usefulness of WebCT, Attitudes towards WebCT, and Perceived Ease of Use was explained very well, considering the reported R^2's in Table 4. Of the variation in scores for Perceived Ease of Use, approximately 24% can be explained by Subjective Norms and the psychology students' Self-Efficacy ratings. A correlation of .38 was found between exogenous variables, Subjective Norms and Self-Efficacy.

The results for Model Two in Table 5 are comparable to those attained for Model One.

Table 5. *Manifest Variable Equations with Estimates for Model Two fitted to Psychology Student Data* [a]

```
PAU21   = 0.00535*PAT_T1  + 1.0000 e1
Std Err   0.00916 GA1
t Value   0.5837

PAU22   = 0.0122*PAT_T1  + 1.0000 e2
Std Err   0.0116 GA2
t Value   1.0530

PGD     = 0.0172*PAT_T1  + 1.0000 e3
Std Err   0.00750 GA3
t Value   2.2959

PAT_T1  = 0.4063*PPU_T1  + 0.0993*PPEU_T1  + 0.00745*PSE_T1 + 0.1707*PSN_T1  + 1.0000 e4
Std Err   0.0450 GA4       0.0491 GA6        0.0146 GA7       0.0763 GA8
t Value   9.0200           2.0236            0.5100           2.2384

PPU_T1  = 0.5680*PPEU_T1 + 1.0000 e5
Std Err   0.0536 GA5
t Value   10.6048
```

[a] PPU_T1=Perceived Usefulness; PPEU_T1=Perceived Ease of Use; PAT_T1=Attitude; PGD=Grades PSE_T1=Self-efficacy; PSN_T1=Subjective Norms; PAU21=Frequency of Use; PAU22=Intensity of Use

Again, Attitudes towards WebCT use predicts neither Frequency of Use nor Intensity of Use, and WebCT Attitudes predict student final grades to the same degree. Perceived Ease of Use, along with Subjective Norms and Self-Efficacy, explained a noteworthy proportion of the variation in scores for Perceived Usefulness of WebCT and Attitudes towards WebCT, as reported in Table 6.

Table 6. *Squared Multiple Correlations for Model Two fitted to Psychology Student Data* [a]

Variable	Error Variance	Total Variance	R-Square
PAU21	0.49425	0.49498	0.00149
PAU22	0.78709	0.79090	0.00482
PGD	0.33118	0.33881	0.0225
PAT_T1	14.12142	25.72644	0.4511
PPU_T1	30.38769	45.31100	0.3294

[a] PPU_T1=Perceived Usefulness; PAT_T1=Attitude; PGD=Grades; PSE_T1=Self-Efficacy; PSN_T1=Subjective Norms; PAU21=Frequency of Use; PAU22=Intensity of Use

Moreover, the correlations between Perceived Ease of Use and the other two exogenous variables were not negligible, with Perceived Ease of Use correlating higher with Self-Efficacy ($r = .48$) than Subjective Norms ($r = .28$). In the end, the evidence tilts in favor of Model Two over Model One, if stringent fit criteria are used. The coefficients associated with Perceived Ease of Use are as viable in Model One as they are in the Model Two. Please note that correlations between Perceived Ease of Use and Self-Efficacy, Perceived Ease of Use and Subjective Norms, and Self-Efficacy and Subjective Norms are .48, .28, and .38 respectively.

Research Question Two: To what extent does the psychology class differ from the engineering class with respect to the factor covariance structures involved in the study?

To answer this question, a multi-sample analysis of the covariance structure differences between psychology and engineering students was undertaken on a variable level, with attention given to each scale considered previously in this study: Perceived Usefulness, Perceived Ease of Use, Attitudes towards WebCT, Subjective Norms, and Self-Efficacy. When a Multi-sample model fits the data well, the implication is that the groups under comparison have very similar covariance structures underlying the patterns in their responses because such models require path coefficients for one group to be equal to path coefficients for the other group. In other words, each multi-sample model constrains the path coefficients for one group to be equal to path coefficients for the other group.

The fit results in Table 7 suggest that covariance structure differences between Engineering and Psychology students are particularly notable with respect to how they responded to questions on the Self-Efficacy and Subjective Norms scales, as neither Multi-sample model fit well.

Table 7. *Fit results for Rival Models by Student Group* [a]

Fit Index	Value
Perceived Usefulness	
Bentler's Comparative Fit Index (CFI)	0.9342
Standardized Root Mean Square Residual (SRMR)	0.0485
Chi-Square	174.9573
Chi-Square DF	60
Pr > Chi-Square	<.0001
Perceived Ease of Use	
Bentler's Comparative Fit Index (CFI)	0.9569
Standardized Root Mean Square Residual (SRMR)	0.0288
Chi-Square	165.9631
Chi-Square DF	60
Pr > Chi-Square	<.0001
Attitude	
Bentler's Comparative Fit Index (CFI)	0.9976
Standardized Root Mean Square Residual (SRMR)	0.0255
Chi-Square	44.0726
Chi-Square DF	40
Pr > Chi-Square	0.3034
Self-Efficacy	
Bentler's Comparative Fit Index (CFI)	0.6153
Standardized Root Mean Square Residual (SRMR)	0.0988
Chi-Square	6799.5032
Chi-Square DF	1404
Pr > Chi-Square	<.0001
Subjective Norms	
Bentler's Comparative Fit Index (CFI)	0.8235
Standardized Root Mean Square Residual (SRMR)	0.0670
Chi-Square	67.4634
Chi-Square DF	24
Pr > Chi-Square	<.0001

[a] Covariance Structure Analysis: Maximum Likelihood Estimation

The manner in which Engineering and Psychology students responded to questions concerning Attitudes towards WebCT are, on the other hand, very similar, as suggested by the a CFI of .99 and a SRMR of .0255. Similar results were found for Perceived Usefulness and Perceived Ease of Use. This may explain why neither Model One nor Two fit the Engineering covariance data well. Perhaps the way in which the Engineering students responded to the Self-Efficacy and Subjective Norms scales was antithetical to the specification of Models One and Two.

Conclusions

Many resources have been invested in the adoption of WebCT. The intent of this study is to assist UCF in offering an alternative educational medium and tailor customized instruction for the purpose of better suiting the wide variety of UCF students with diverse backgrounds. Due to limited financial and human resources, it is incumbent upon the university to make learned decisions when implementing an information infrastructure project. A smart marketing approach can take full advantage of money spent on this initiative. The results presented in this current study suggest that extended adaptations of the Technology Acceptance Model are not as suitable for Engineering students as they are for Psychology students. Moreover, although both competing models hypothesized are capable of explaining the attitudes of students taking a psychology course, neither ultimately does a very good job of explaining how frequently or how long students make use of WebCT while completing their coursework. Still more, neither model can predict very well the Final Grades that students earn. In the end, with respect to WebCT use, final grades were not associated with the Frequency of WebCT use ($r = -0.0647$, $p = .3290$), and very little with the Intensity of WebCT (i.e., duration of WebCT use per session; $r = -.1690$, $p = .0102$). Indeed, the ever slight association existing between Intensity and student grades would suggest that students who used WebCT for longer periods, tended to do slightly less well.

This research only applies to one university setting. Its results can only be applicable to its similar settings. Because the sample was selected purposively, the study can be skewed to some extent. The Technology Acceptance Model was not adapted well by the engineering class, suggesting that one model does not fit all. New or modified TAM models are in demand. Additionally, two competing models failed to predict to a satisfactory degree the three outcome variables: Frequency of Use, Intensity of Use, and Grades, which implies that more external variables are needed.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84(2), 191-215.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions, and behavioral impacts. *International Journal of Man Machine Studies, 38,* 475-487.

Dziuban, C., & Moskal, P. (2001). Evaluating distributed learning in metropolitan universities. *Metropolitan Universities, 12*(1), 41-49.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research.* Reading, MA: Addison-Wesley.

Harris, R. W. (1999). Attitudes towards end-user computing: A structural equation model. *Behaviour & Information Technology, 18*(2), 109-125.

Hu, L., & Bentler, P. M. (1999). Cut criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Hubona, G. S., & Blanton, J. E. (1996). Evaluating system design features. *International Journal of Human-Computer Studies, 44*, 93-118.

Igbaria, M., Zinatelli, N., Cragg, P., & Cavaye, A. L. M. (1997). Personal computing acceptance factors in small firms: A structural equation model. *MIS Quarterly, 21*, 279-305.

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management, 40*, 191-204.

Lu, H. P., Yu, H. J., & Lu, S. S. K. (2001). The effects of cognitive style and model type on DSS acceptance: An empirical study. *European Journal of Operational Research, 131*, 649-663.

Marcoulides, G., & Scott L. Hershberger, S. (1997). *Multivariate statistical methods: A first course.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Moskal, P. D., & Dziuban, C. D. (2001). Present and future directions for assessing cybereducation: The changing research paradigm. In L. R. Vandervert, L. V. Shavinina, & R. A. Cornell (Eds.), *Cybereducation: The future of long distance learning* (pp. 157-184). Larchmont, NY: Many Ann

Liebert.

Pan, C. (2003). *System use of WebCT in the light of the Technology Acceptance Model: A student perspective.* Unpublished doctoral dissertation, University of Central Florida, Orlando.

Pan, C., Sivo, S., & Brophy, J. (in press). Students' attitude in a Web-enhanced hybrid course: A structural equation modeling inquiry. *Journal of Educational Media & Library Sciences.*

Pan, C., Sivo, S., Gunter, G., & Cornell, R. (under review). Predictive variables in the Technology Acceptance Model: A multivariate method. *Journal of Research on Technology in Education.*

Robinson, L., Jr. (2001). *Sales force use of technology: An extension of the technology acceptance model, including antecedents and outcomes relevant in professional selling organizations.* Unpublished doctoral dissertation, University of South Florida, Tampa.

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7^{th} ed.). New York: McGraw-Hill.

Sivo, S., Pan, C., & Brophy, J. (in press). Temporal cross-lagged effects between subjective norms and students' attitudes regarding the use of technology. *Journal of Educational Media & Library Sciences.*

Venkatesh, V., & Davis, F. D. (1994). Modeling the determinants of perceived ease of use. *Proceedings of the 15^{th} International Conference on Information Systems, Vancouver, British Columbia, Canada*, 213-228.

Wolski, S., & Jackson, S. (1999, February). *Technology diffusion within educational institutions: applying the technology acceptance model.* Paper presented at the 10^{th} Society for Information Technology & Teacher Education International Conference, San Antonio, TX.

Creative Instruction: Case of Indonesiana and Filipiniana

Paulina Pannen
Jamaludin
Universitas Terbuka
Indonesia

Abstract

This paper describes the results of a study on creative instruction of science and mathematics in Indonesia and in the Philippines. The study focuses on the perception and consideration of teachers in the design process, and the experience of creative instruction using arts and culture, i.e., Indonesiana and Filipiniana. The use of Indonesiana/Filipiniana in science and mathematics instruction is assumed to create a creative, lively, and joyful instruction and learning process, which is highly conducive for students' understanding

Creative instruction using Indonesiana and Filipiniana is aimed at improving the quality of instruction in science and mathematics. In being creative, teachers are challenged to make a difference in their instruction and learning via creative integration of Indonesiana and Filipiniana into the instruction of science and mathematics. This creative teaching differs from what can be seen in school at present, where science and mathematics instruction is viewed as finding correct answers.

The study employed several data collection strategies, i.e., introductory seminar, workshop, review and revision, implementation (try-out), and survey. The respondents were 25 teachers, 18 from Indonesia and 13 from the Philippines, and their students in their respective class. The results of the study indicate that the creative instruction of science and mathematics using Indonesiana and Filipiniana was perceived positive by students and teachers. During the implementation, teachers' and students' expectations on having an engaging, exciting, and rewarding instruction can be fulfilled. The well-designed lesson plan, the participation of the students, and the availability of Indonesiana/Filipiniana have been the factors contributing to the successful instruction. However, the study found that teachers considered the design process tedious and difficult, they did not have good knowledge on their own arts and cultures, their orientation was mostly one-way teacher and content oriented, and they were bound to follow the "tradition".

It is recommended that design process be an essential part of the teachers' instructional ritual at schools. Therefore, further efforts in familiarizing teachers with a design process must be taken to enable the teachers to gain deeper understanding of their topic, the local arts and culture related to their topics. It is also recommended that the design process takes the team approach, so that each member of the team can assist one another in different aspects of design. In addition, a new culture of instruction must also be fostered in order to enable the teachers to be creative, instead of just following the tradition.

Introduction

Despites of various efforts to conserve traditional arts and culture in Indonesia and the Philippines, both Indonesiana and Filipiniana thus far are seen as a separate entity outside the formal schooling system, and rarely counted as the contextual surroundings which play a significant role in the development of students understanding in schools. Compared to science and mathematics, Fajardo (2002) asserts that *"... science and mathematics are dominating, in the sense that teachers, etc. see them as up there in the hierarchy of learning. The arts and humanities have always been seen as less than that ...".*

Creative integration of Indonesiana and Filipiniana into science and mathematics teaching is a culture-based approach which believes that learning is very closely connected to one's culture (Jarvis, Holford, Griffin, 1998. What count as knowledge differs between cultural contexts, and this view implies, according to Barret (2002), that teachers should consider using culture-based methodologies in the classroom. This culture-based approach is expected to enable teacher to create a joyful, meaningful, and contextual science and mathematics teaching and learning environment, and to integrate both scientific values of science and mathematics and beautiful images of arts and culture together in one learning experience.

Creative instruction using Indonesiana and Filipiniana is aimed at improving the quality of instruction in science and mathematics. In being creative, teachers are challenged to make a difference in their instruction and learning from what can be seen in school at present \, through creative integration of

Indonesiana and Filipiniana into the instruction of science and mathematics. This study is about creative instruction of science and mathematics through the use of Indonesiana and Filipiniana. It focuses on the perception and consideration of teachers in the design process, and the experience of creative instruction in science and mathematics using arts and culture, i.e., Indonesiana and Filipiniana.

Previous Studies

The infusion of arts and culture across curriculum has been widely studied. The works of the Artsvision (Eisner, 1994; Greene, 1995, Upitis, 1997) have resulted in a compelling recommendation to use the arts not only as a discipline, but also as a means of teaching across the curriculum. Much have also been written about the value of the arts and culture as a tool for the teaching and learning of many skills and concepts beyond the arts themselves (Elster, 2001; Grauer, et al., 2001). Goldberg (2001) also confirms that arts and culture are powerful tools to motivate students to apply their knowledge, work cooperatively, and make connections across content areas.

In the Philippines, Rapanut, et al., (1996) employed the ethno mathematics model to study the indigenous mathematical concepts operating in various aspects of Kankana-ey culture in Mountain Province, specifically weaving patterns, gong music, and kinship system. The study resulted in four lesson plans of mathematics for first year high-school. Other efforts were also taken by NISMED through its publications which infuse arts and culture into science and mathematics education in the Philippines (NISMED, 2001).

In Indonesia, in 1999, Universitas Negeri Semarang introduced the so-called society, education, and technology (SETS) – an approach that takes community and society into consideration for teaching of subject matter, especially science and technology. Later, Pekerti (2001) introduced the teaching of mathematics through music in elementary schools, and Riyanto (2002) also introduced the local games in several areas of Indonesia for learning mathematics. Another effort includes the introduction of chemistry teaching through cooking.

Creative Instruction

Creativity is probably one of the most important aspects of education. According to Baer (1997), c*reativity refers to anything someone does in a way that is original to the creator and that is appropriate to the purpose or goal of the creator*. While Gardner (1993) defines that *a creative individual is a person who regularly solves problems, fashions products, or defines new questions in a domain in a way that is initially considered novel but that ultimately becomes accepted in a particular setting*. Thus, a creative person considers and even questions the actions s/he performs, and the reflective nature of a creative person sets her aside from the person who mechanically carries out an assignment or follows through on a task. Goldberg (2001) states that *being creative is essentially a process of awareness that can open new doors and change existing perceptions.* In this case, creativity is not limited to a few special, highly valued artistic or scientific activities, nor is it limited to only a few outstanding ideas or works of art conceived by famous people. It is a variety of anything, the kind all of human being have, use, enjoy, and has meaning to the lives because it makes a difference.

Being creative means making a difference or making changes to the existing ones, therefore, creative instruction is different from the existing picture of instruction. In being creative, teachers are challenged to create an enjoyful and meaningful instruction through the use of different strategies which make a difference in their instruction. In this case, simply being original, or just to follow the order (or the tradition), or to mechanically carry out an ordinary teaching is not enough. Sometimes, being different, changing the existing perceptions, or even "considerably" weird is expected in being creative. At the end, this creative instruction is expected to bring about improvement in the instructional process, through a different picture of instructional process.

For many students, success in school has very little to do with true understanding, and much to do with coverage of the curriculum. Curriculum is held absolute and teachers are reticent to tamper with it even when students are clearly not understanding important concepts. Even students who are capable of demonstrating success, who pass tests with high marks and obtain honors diplomas, frequently do not connect the information they receive in school to interpretations of the world around them. In creative instruction, teachers follow various ways to create learning environments where they and their students will be able to work together as active learners constructing knowledge. Through such a process, they are able to invite students to experience the world's richness, empower them to ask their questions and seek their own answers, and challenge them to understand the world's complexities. Furthermore, they will also be able to creatively teaching any body of knowledge through scientific inquiry based on the local context and local problems, to be applied in the local cultural community where the students come from. Then, the classroom will not be dominated by the teacher talks, teachers will not transfer knowledge and expect students to identify and replicate the information (or give one correct answer), teachers will not rely heavily on (a single) textbook nor presenting one set of truth, students are challenged into working cooperatively and collaboratively in solving complex issues, and students opinion (and argumentation) in many forms of representations are expected. The success of the school in guiding the students into their development depends ultimately on the teachers – their competence, values, and commitment, including being creative in the making of meaningful instructional process, a different instruction.

There are many ways for teachers to make a difference in their instruction. Teachers can use different approaches, different materials and media for learning, new sequence of instruction, alternative assessment style, and build a different learning environment. Being creative teacher of science and mathematics means using a variety of instructional strategies in the repertoire to create enjoyful and meaningful instructional process. The use of Indonesia and Filipiniana in science and mathematics instruction is one creative attempt involving different instructional techniques than the common topic-based teaching techniques. It suggests that an instructional process is not for the mere transfer of teachers' knowledge or information from a textbook to students, or for the completion of prescribed topics in the curriculum, but it is for integrated, contextualized, and meaningful understanding – that the body of knowledge gain in school will have place in the students own life and cultural environment

Indonesiana and Filipiniana

According to E.B. Taylor (1871), culture is that complex whole which includes knowledge, belief, art, law, morals, custom, and any other capabilities and habits acquired by man as a member of society. A cultural work is created through creative and artistic process, appreciation, and also learning process which are based on the cultural phenomenon.

Acquaintance with culture and the works of art plays a fundamental role and offers a rewarding experience. According to Fajardo and Flores (2002), culture and the works of art introduce non-algorithmic reasoning, complex thinking, thinking that yields multiple rather than unique solutions, multiple criteria =, which sometimes in conflict with each other, thinking that is laced with uncertainty, self-regulation of the thinking processitself, strategies to impose meanings and finding structure in apparent disorder, and nuance judgment and interpretation. It is through this rewarding experience that cultural values will be appreciated and internalized.

Goldberg (2001) asserts that culture and arts are powerful and natural tools for everyday learning. The integration of culture and arts provide strategic methods for motivating students to apply their knowledge, work cooperatively, and make connections across content areas. It challenges teachers and educators to be able to collaborate with artists and students, and to be creative in their teaching. Understanding culture and arts, one learn to think and work across traditional disciplines since culture and arts integrate knowledge from various branches. It encourages the suppleness of mind, a toleration for ambiguity, a taste for nuance, and the ability to made trade-off among alternative courses of action. It is through culture and arts, that one builds understanding of diversity, multicultural dimensions of world, cooperative work, and value content (quality of work and high achievement). The exploration and interaction with culture and art and its values accommodate the process of building intellectual structures as well as affective capacities that continuously change, interact, and combine, to promote the development of human understanding (Goldberg, 2001).

Indonesiana and Filipiniana are the terms loosely used to describe the culture and arts of the Indonesian and the Philippines. Indonesiana and Filipiniana serve as context where learning take place, i.e., where students are engaged in intellectual discourses with information of subject matters, where scientific concepts, principles, and theories are learned, where students apply those scientific concepts, principles, and theories to solve their lives' problems in their community, and where students creatively develop those scientific concepts, principles, and theories for further development of the subject matter as well as the well-being of the community. In that case, Indonesiana and Filipiniana provide a complex whole of knowledge, belief, art, law, morals, custom, and any other capabilities and habits acquired by man as a member of society. The new scientific information presented to students is to be built based on the existing scheme in their mind, i.e., their cultural knowledge and experience of Indonesiana or Filipiniana.

Despites its richness, thus far, Indonesiana and Filipiniana are not much regarded as part of the instructional process nor counted as the contextual surroundings. The learning process of any subject matters in schools are detached from its context – Indonesiana and Filipiniana. A number of studies indicate that students do not develop an understanding of science and mathematics that are useful for their lives in a socio-cultural community. Most students memorize science terms and mathematical theorems without understanding its applications in their cultural community. In the practices of science and mathematics instruction, teachers are too focused on teaching for the sake of science and mathematics as fields of knowledge (Tate IV, 2001; Lythcott and Stewart, 2001). Teaching efforts are mostly focused in transferring science and mathematics knowledge to students as prescribed by curriculum, which has led the teaching process into a relatively dull and dry process, and a limited classroom effort. Such a situation has contributed to the fact that students do not comprehend the role of science and mathematics, or any other subject matters, in their own lives, nor they appreciate the learning of science and mathematics in schools.

According to Gardner, Brooks and Brooks (1993) and Vigotsky's work on constructivism, contextualizing subject matter into local needs, problems, and context, will help students to construct better scheme of knowledge than loose decontextualized subject matter. Shepard (2002) also claims that new learning is shaped by prior knowledge and cultural perspectives. Therefore, the integration of Indonesiana and Filipiniana in science and mathematics instruction is necessary to address learning from local cultural and surroundings, and the application of science and mathematics to local problems and needs within the students' cultural community. It is a creative alternative to create fun and interest in learning of science and mathematics, and also to achieve better learning results.

Assumptions

The use of Indonesiana and Filipiniana in science and mathematics instruction embraces tightly the notion of centricity: students' culturally relevant knowledge and experiences are the center of learning. Indonesiana and Filipiniana is the cultural community where the students belong to, and the context where knowledge is learned, developed, and applied. Prior knowledge, understanding, beliefs, and attitudes of the students that are translated as cultural relevant knowledge, understanding, beliefs and attitudes held firm by the students' cultural community provide a springboard for teachers to introduce new information of science and mathematics (Tate IV, 2001; Lythcott and Stewart, 2001). The richness of Indonesiana and Filipiniana provides a context where students make connections and find metaphors of their new information to their real situation. Furthermore, it serves as a place that can be benefited from the students knowledge of science and mathematics, i.e., through the application of scientific and mathematical knowledge for development of the cultural community. Cultural community where conflicts and problems arise, and solutions are seek should be the center or the preeminent place for every learning endeavor to embark.

Thus, within the framework of integrating Indonesiana and Filipiniana into science and mathematics instruction, the instruction was expected to be designed creatively by taking into account the students relevant cultural knowledge as the springboard to introduce new information, and the integrative use of arts, materials, and culture. That way, it was expected that students will find science and mathematics meaningful in their own cultural life, and they will find arts and culture meaningful in the science and mathematics they learn from school.

Methodology

This study was conducted in the Philippines, i.e., Los Banos area, and Iloilo (Visayas) for 6 months during April – October 2002; and in Indonesia, i.e., Pondok Cabe, Jakarta area for 5 months during January – May 2003. The operational framework of this study is as follows:

Stages of learning	**Functions and Coverage of Indonesiana and Filipiniana**	
	Learning with Indonesiana and Filipiniana	*Learning Through Culture, Indigenous Arts and Materials*
Knowledge acquisition		
Problem Solving Skills		
Scientific Inquiry Skills		
	In Elementary and Secondary – Science and Mathematics	

To study the integration and use of Indonesiana and Filipiniana in the teaching environment, a series of methods and strategies were employed, i.e., seminar and workshop, classroom observation for implementation, survey and observation.

Survey and Observation

Survey and observation I was conducted prior to the introduction of the use of Indonesiana and Filipiniana in science and mathematics instruction. It was intended to capture the existing common instructional practices of science and mathematics at schools, in addition, to tapping students and teachers' perceptions on science and mathematics, on the tradition of teaching and learning of science and mathematics, and on the use of culture and arts for teaching and learning of science and mathematics. Survey and observation II was conducted during the try out (implementation) of the integration of Indonesiana and Filipiniana in science and mathematics instruction. It was intended to tap teachers and students' perceptions on the experience of teaching and learning using Indonesiana and Filipiniana.

Seminar and Workshop for Planning

An introductory seminar was conducted to introduce the idea of science and mathematics instruction through Indonesiana and Filipiniana to participating teachers. Following the seminar, a workshop on planning of creative science and mathematics instruction through Indonesiana and Filipiniana was carried out. The workshop was meant to provide a supervised and facilitated opportunity for teachers to develop their lesson plans on science and mathematics instruction using Indonesiana and Filipiniana. The facilitation process carried out by faculty members was continued after the workshop through a series of less formal meetings and individual tutorials.

Implementation

In Indonesia, 22 teachers participated, while in the Philippines, 12 teachers participated. The decision to participate was made by the individual teacher, based on the time consideration and self-confident. Before the implementation, preparation of teachers were monitored via telephone or personal visit, and problems were noted. During the implementation, narrative notes were taken to record the process of the teaching and learning. In addition, digital images of the implementation were also taken as supporting information.

Data collected in this study were qualitative data of written perception (anecdotal records), images, and written products (lesson plans). Therefore, for the purpose of data analysis, skimming, classification, concept mapping, and summing up techniques were being used, in addition to anecdotal quotations.

Results

Initial Perception

The initial teachers' perception on the integration of Indonesiana and Filipiniana in science and mathematics instruction were doubtful, even negative. Some Filipino teachers claimed that science and mathematics instruction through Filipiniana will *make the class/session interesting and more relevant, it will enhance the ability of teachers, the students will find it more meaningful, it is a form of interdisciplinary learning process, and it is worth trying*. One teacher was reporting that, *"... it should. Children learn faster and appreciate more what they learn if they are involved with materials that are familiar to them which they see more in their environment"*. Nevertheless, some teachers indicated that they *do not know the idea,* and perhaps *it will be difficult at first*. Indonesian teachers, however, provided more skeptical and indecisive comments: "… *it may create a more interesting session, can be attractive to students, perhaps give more variation in teaching, and, we have to be creative…"*.

The Filipino teachers also expressed that the concept should be *supported by administration, teachers should be watchful for the misconceptions, and teachers should have sufficient knowledge*. This view was confirmed by the Indonesian teachers, who specifically mentioned that supports from the school's principals, and also from faculty member of the nearby university (i.e., faculty of education) are needed. These expressions indicated a relatively unsecured feeling of teachers to embark on a new venture. Administration is known to be the major obstacle for changes in the tradition of teaching. In addition, teachers' mastery of subject matter as well as Indonesiana and Filipiniana are also important consideration. Therefore, the Indonesian teachers specifically perceived that such a situation needs facilitation from faculty members (and also experts on Indonesiana) from the nearby university.

Students' perception, however, are more varied. Some Indonesian and Filipino students expressed positive perception that it *offers limitless opportunities for learning, learning becomes meaningful, art gives color to the subject, will help bridge the gap in learning, courses will be easier, unique, lively, interactive, fun, challenging, and effective*. One student said that, *".. I would really like that idea… help to bridge the gap* *between learning and fun"*. However, there were some students who considered the idea to be *complicated, two different things – one deals with numbers and the other deals with creativity, hard to do, confusing, and unthinkable*. According to one student, *"I think it is quite impossible to combine science and mathematics and Filipiniana, because these are 2 different things, it will just lead to confusion"*. An Indonesian student claimed that "*it will be boring, chaos, and will not match"*. Still some students were doubtful: "*I do not expect too much*". These students' perception provided a challenge to teachers.

The Design And Preparation Process

Each teacher are expected to design 2 lessons in science and mathematics at any level/grade for implementation. The design framework is as follows:

Instructional Activities		**Constructivism (Brooks & Brooks)**
Engagement	To focus students thinking on the current concepts or skill to be learn, to connect between students past experiences and the present lesson	Introduction
Exploration	To develop students understanding (and discovering) a concept by engaging in concrete experience with materials, by exploring on their own	• Authentic Task • Active learning engagement with phenomena/problems of emerging relevance to students • Use and application of knowledge based on primary concepts and students' point of view • Use of learning communities
Explanation	To introduce formal vocabulary or students verbalize understanding about the explorations they have been engaged	
Elaboration	For students to gain deeper understanding of the concept by engaging in additional activities	
Evaluation	To assess students understanding and their knowledge	Multiple Representations

During the design and preparation process, the teachers were facilitated by humanity and arts experts, instructional designers, and subject matter experts (science and mathematics educators). To start with, there were many assumptions made regarding the teachers competence and mastery on subject matter area, their knowledge and skills in lesson planning, their knowledge in Indonesiana and Filipiniana, and their motivation to be creative, and their level of creativity. It was found out that not all of the assumptions were correct, thus called for alternative strategy in the design and preparation process.

Teachers mastery on subject matter area was not as high as expected. Although most teachers in Indonesia have their Sarjana degree, and in the Philippines have their Bachelors and some have their master degree, their mastery on subject matter area is somewhat discouraging. Especially, they master the subject matter as content of a field of discipline, not within a context of a subject matter to be taught (instructional content) to students in the elementary as well as secondary schools. The teachers indicated vague knowledge of what goes beyond the subject matter – the reasoning and rationale of each concepts, principles or theorems, the strategy and techniques to easily learn and master the concepts, principles, or theorems, the structure of the knowledge in science and mathematics, and also the essence of learning science and mathematics. Their mastery of the subject matter is more on the surface level, it does not go into deep meaningful level of understanding. In other words, they did not obtain the cognitive strategies and higher order learning skills when they master the subject matter, they did not realize why students should master the subject matter and what competencies should be achieved when students study the subject matter. For this reason, teachers faced difficulties in matching the subject matter to culture and arts concept. It was relatively difficult for them to analyze the analogical and metaphorical connection between science and mathematics and culture and arts.

The situation was worsened since the teachers' knowledge on Indonesiana or Filipiniana were highly limited. The study found that teachers did not have good knowledge on their own local arts and culture. It took the teachers a journey through a stack of the *Indonesian Heritage Encyclopedia* and the *CCP Encyclopedia of Philippine Arts*, also a series of discussion with the humanity and art experts, for them to be able to select one piece of Indonesiana or Filipiniana to be integrated into their instruction. This situation was contradictory to the survey carried out earlier, in which teachers were very determined in providing variety of examples of Indonesiana or Filipiniana to be integrated in their instruction. Their determination did not show up during the design and preparation process. Some teachers were seemed lost as they were trying to tie their subject with specific piece of Indonesiana or Filipiniana. This perhaps is due to superficial understanding of their subject – the structure and essence of their subject, or merely lack of knowledge on Indonesiana or Filipiniana.

It was very surprising to find that lesson planning was a real burden for teachers. Teachers reported that in practice, design process was not part of their instruction ritual. The design process was considered tedious and difficult, on top of their overload teaching assignment. Indonesian teachers said that they usually did not need to write their own lesson plan, since the national curriculum has provided everything for them, including the lesson plan for each topic of their subject. The Filipino teachers claimed the same situation, and added that they were not required to write lesson plan to be able to teach. It was a fact that lesson planning was not a part of the teaching ritual at schools, nor the teachers habit. During the design and preparation process, the lesson planning activities were perceived hard and difficult by the teachers, and it was a real challenge for the instructional designers during the facilitation activity.

Another discouraging issue was the teachers' motivation to be creative and their level of creativity. Both Indonesian and Filipino teachers were very conformist and uniformist. their orientation was mostly one-way-teacher and content-oriented, instead of student-oriented instruction, and they were bound to follow the "tradition", instead of trying new strategies. They were very careful not to go out of the tradition, afraid of being different from the present practice or from anybody else. This situation was tapped from the first draft of plans they wrote, which indicated high similarity with the common picture of instruction where the teachers are dominant, students are passive, interaction is limited, correct answer is demanded. The new culture-based approach which introduces the concept of joyful and meaningful learning taken basically from constructivism was perceived to be out of the tradition.

It was found out that the four assumptions underlying the design and preparation process were not applicable. The teachers perceived that the design and preparation process demanded a lot of effort ("*it was a lot of works and preparation, you have to know your subject well to select any piece of arts, planning is tedious and difficult*"), since the teachers had to write their own plan while not having enough knowledge about Indonesiana or Filipiniana, or had difficulties in finding pieces of Indonesiana or Filipiniana to fit their plan. The teachers also indicated that they need more assistance from subject matter experts in analyzing specific concepts of science and mathematics in regard to the pieces of Indonesiana or Filipiniana to be integrated. Based on the situation, the design and preparation process was extended to go beyond the two-day workshop into one full month of intensive informal meetings and discussion for the Indonesian teachers and six weeks for the Filipino teachers.

Perception on the Experience

During the try-out, classes were no longer dominated by teachers anymore. Although teachers were still leading the students through challenging activities for higher order thinking, students were relatively given freedom to express their own opinion, and to have different perspectives based on sound justification. Assessment came in various forms, tests as well as non-tests, where both teachers and students become the assessors. The use of media – Indonesiana or Filipiniana, material, laboratory equipment, posters, audiovisual material, realia and models – and concrete examples were intensive and varied. Activities were varied, challenging, mindful, and meaningful both to the topics and to the students. The teachers were relatively ready and prepared with their session. Student's participation was very high. Instructional sequence in the form of worksheet and student's activity sheet, were provided for students to be able to participate in the process independently.

The try out demonstrated that teachers have tried to put each steps taken and each concept into the context of students' daily life. Authentic tasks and ample of activities were given to allow intensive and joyful interaction between teachers and students. Contextual use and application of knowledge were indicated – through examples given, the pieces of Indonesiana or Filipiniana being integrated, and also through extended examples provided by the students. In Indonesia, instructional activities were still focused in schools, while in the Philippines, the use of learning communities was tried out, i.e., public market, craft centers, museums, botanical garden. Multiple representation of understanding was also implemented, i.e., test, poems, drama/role play, product, model. Indeed, the picture has been a successful steps taken by the teachers to change their instructional paradigm.

From the perspective of infusion of Indonesiana and Filipiniana, the coverage of Indonesiana and Filipiniana was expanded to include the indigenous arts, culture, and indigenous materials (flora, fauna, places, etc.) found in the local area. It is not limited to specific art or material that is native to the people, but more to the arts and materials that are locally available and meaningful to the people's life. This is to give a way to teachers who do not have adequate knowledge about Indonesiana or Filipiniana.

The teacher perceived the try-out as a new experience, a challenging and a fulfilling one. Some teachers were satisfied with their implementation, especially when they knew that students really learned from the try-out sessions ("*posttest results were good*"), the students had fun and enjoyed it, the students are active, free in expressing themselves, more motivated, and more tolerant to others, and the teacher proved they can do it ("*I can do it*"). One teacher said that the experience *"…is a complete revision of my usual approach to the students".* There are some Indonesian teachers who expressed their intention to revise their plan, and or to write another lesson plan and conduct another try out to improve the results of their endeavor.

The try out told the teachers successful stories, therefore, they acknowledged that it needs to be followed up, so they become more familiar with the idea and more skillful. They noticed that their plan could serve important guidance during the implementation process, and that the try-out has provided them with a stimulus to change, be more creative, see other possibilities in teaching, and widen their knowledge of the subject as well as of their own culture and arts. Nevertheless, impeding factors such as poor knowledge of Indonesiana and Filipiniana, lack of design skills, and new skills to be aquired in integrating Indonesiana and Filipiniana in their teaching of science and mathematics had caused the creative process considerably difficult for teachers, even when successful stories are experienced.

The teachers perceived that students' eagerness and willingness to participate has contributed to the successful implementation, in addition to a well-designed plan, teachers' readiness, and the availability of Indonesiana and Filipiniana. On the other side, students' perception on their experience are encouraging, i.e., they enjoyed the experience, and they would like some more of the "*cool*" sessions which were *"fun and learning at the same time*", *teach many things*", allow them to make more friends, use their creative ideas, learn the topics easier, learn to cooperate and that the activity enable them to interact more with their teachers.

Remarks

The the idea of using Indonesiana and Filipiniana in science and mathematics teaching is calling for an instructional paradigm shift to a students-centered approach which is the opposite of the current practice – the teacher-centered approach. The idea also calls for multiple representations of students understanding, while in the present practice, (paper and pencil) test (and numerical score) is still required as a proof of students' achievement.

In this study, resistance toward the idea had been moderate. Some negative perception from teachers existed, especially in the "*tedious and difficult*" planning process, and observation process which made the teachers nervous. However, teachers acknowledged that a well-designed lesson plan was needed for a creative instruction using Indonesiana/Filipiniana. Therefore, design process must be made an essential part of the teachers' instructional ritual at schools. It maybe tedious and difficult at first, however, it enables the teachers to gain deeper understanding of their topic, the local arts and culture related to their topics when they design their lessons. Further efforts in familiarizing teachers with a design process must be taken. It is also recommended that the design process takes the team approach, so that each member of the team can assist one another in different aspects of design. In addition, a new culture of instruction must also be fostered in order to enable the teachers to be creative, instead of just following the tradition.

Willingness of the teachers to participate because the idea was novel to them was encouraging. In addition, there were cases, when culture and art were used/integrated, they were used in a trivial way, for example to start a lesson, or for the sake of students being active. Therefore, meaningful and contextual uses of Indonesiana and Filipiniana need to be encouraged through further implementation and studies. Exploration and studies on wider variety of Indonesiana and Filipiniana for integration across curriculum are also needed.

The try out has shown a different picture of teaching and learning process, i.e., the enjoyful and exciting lesson as opposed to a difficult and boring lesson. Both teachers and students expectation were fulfilled. It has been an important successful small step for teachers to embark on the successive steps of changes in teaching paradigm, and to maintain their motivation on engaging in the making of different and meaningful instruction.

The underlying premise of integrating the Filipiniana into the teaching of science and mathematics is the application of science and mathematics in relation to real world problems in the context of everyday life. In addition to making the Filipiniana known to the Philippines' children, such an integration also supports the Philippines' government efforts to conserve and preserve the Filipiniana, as their national identity. Furthermore, it also opens a way for the Philippines' children to apply their knowledge and skills (learned from school) to develop and to enrich their own arts and culture.

Based on the findings, it is therefore recommended for Indonesian as well as Filipino teachers to have more exposure on Indonesiana and Filipiniana respectively. Further, since the try out has provided positive results, the effort of providing joyful and meaningful instruction via creative means must continued to evolve. Teachers need further guidance and facilitation to be creative and to make a different in their instruction based on a well-designed lesson plan. Further, this study is to serve as a springboard for further endeavor for many individuals and institutions interested in engaging in similar research ventures

References

Barret, D.E. (2002) Does Culture Matter? *Society*, Vol. 40, No. 1, November/December 2002, pp. 92-96.

Bernardo, A.B.I. (2002) Do You Know Where You're Going? Global Trends in Mathematics and Science Education Reform and Their Implications for Philippine Reform Efforts. *UP-CIDS Chronicle*, July-December 1998, 2002.

Brooks, J.G. & Brooks, M.G. (1993) *In Search of Understanding: The Case for Constructivist Classrooms.* Alexandria, VA.: Association for Supervision and Curriculum Development.

D'Ambrosio, U. (1985) Ethnomathematics and Its Place in the History and Pedagogy of Mathematics. *For the Learning of Mathematics*, 5, 44-48.

De Guzman, G. (2002) *Integrating Culture and Art in Science and Mathematics*. Paper presented in Seminar Workshop on Contextual Constructivism: Creative Teaching through Indigenous Arts. UP Visayas, Iloilo, September 7, 2002.

Eisner, E.W. (1994) *Cognition and Curriculum Reconsidered*. New York: Teachers College Press.

Elster, A. (2001) Learning Through the Arts: Program Goals, Features, and Pilot Results. *International Journal of Education and the Arts*. Vol. 2, No. 7, 2001

Fajardo, B.V., Flores, P.D. (a.n.) *HUM II: Art and Society*. Los Banos: UP-OU.

Fajardo, B.V. (2002) *Indigenous Art and Constructing a Contextualized Curriculum*. Paper presented in Seminar Workshop on Contextual Constructivism: Creative Teaching through Indigenous Arts. UP-OU, Los Banos, June 27, 2002.

Fernandez, H.R. (2002) *Towards a Culture-Centered Approach to General Education and Arts Education*. Paper presented in Workshop for Faculty of the Philippine High School for the Arts, Los Banos, July 18, 2002.

Gardner, H. (1993) *Multiple Intelligences: The Theory in Practice*. New York: Basic Books.

Goldberg, M. (2001) *Arts and Learning: An Integrated Approach to Teaching and Learning in Multicultural and Multilingual Settings*. 2nd Ed. New York: Addison Wesley Longman.

Grant, C.A. & Gomez, M.L. (2001) *Campus and Classroom: Making Schooling Multicultural*. 2nd Ed. Upper Saddle River, New Jersey: Merrill – Prentice Hall.

Grauer, K ., et. al. (2001) Images for Understanding: Snapshots of *Learning through the Arts™*. *International Journal of Education and the Arts*, Vol 2, No. 9, November 2001

Lythcott, J. & Stewart, J. (2001) Learning Science as a Transformative Experience. In Grant, C.A. & Gomez, M.L. (2001) *Campus and Classroom: Making Schooling Multicultural*. 2nd Ed. Upper Saddle River, New Jersey: Merrill – Prentice Hall.

National Institute for Science and Mathematics Education Development (2001) *One Hundred Years of Science and Mathematics Education in the Philippines.* Diliman, Quezon City: NISMED and Foundation for the Advancement of Science Education, Inc.

Padolina, M.C.D. (1999) Science Education in the Philippines. 24 February 1999.

Pannen, P. (2002) *Contextual Constructivism: Creative Teaching Through Indigenous Arts.* Research Report – Asia Fellows Program. Bangkok: Asian Scholarship Foundation, Ford Foundation.

Primadi, T. (1998) *Proses Kreasi, Apresiasi*, *Belajar*. Bandung: Penerbit ITB.

Rapanut, T. A. et al. (1996) Algebra of the Weaving Patterns, Music and Kinship System of the Kankana-Ey. Baguio City: UPC-Baguio, The Faculty of the Discipline of Mathematics.

Richardson, V. (1997), Constructivist Teaching and Teacher Education: Theory and Practice. In Richardson, V. ed. (1997) *Constructivist Teacher Education: Building New Understanding*. Washington, D.C.: The Falmer Press.

Tate IV, W. F. (2001) Mathematizing and the Democracy: The Need for an Education That is Multicultural and Social Reconstructionist. In Grant, C.A. & Gomez, M.L. (2001) *Campus and Classroom: Making Schooling Multicultural*. 2nd Ed. Upper Saddle River, New Jersey: Merrill – Prentice Hall.

Upitis, R. (1998) Pre-service and inservice teacher education in the arts: Toward a National Agenda. In B. Roberts (ed.) *Connect, Combine, Communicate: Revitalizing the Arts in Canadian Schools.* Sydney, NSW: University College of Cape Breton Press.

Vygotsky: The social formation of mind in Driscoll, M.P., (2000). *Psychology of Learning for Instruction*. 2nd ed. Boston: Allyn & Bacon

Note: The Philippines portion of this study was funded by the Asian Scholarship Foundation – Asia Fellows Programs – Cohort 3, 2002. The Indonesian portion of this study was funded by DGHE, Ministry of Education, and Universitas Terbuk, Indonesia.

The Effect of Teacher Education Students' Epistemological Beliefs on Program Usage and Learning Outcomes in a Case-Based Hypermedia Learning Environment

Hsinyi Peng
National Chiao Tung University

Abstract

Personal epistemological beliefs, how individuals believe about the nature of knowledge, have been shown to influence learning in many aspects. This presentation examines the effects of epistemological beliefs of teacher education students and their learning in a case-based hypermedia learning environment. Data are collected via a set of questionnaires, pre- and post-knowledge tests, and audit trail analysis of usage and artifacts on problem solving.

Introduction

The current school reforms in teacher education have focused on increasing the opportunities for students to gain more authentic classroom experience via field placement and technology based on real classroom scenarios. Echoing the school reforms, a number of instructional approaches have been implemented, including case-based learning (CBL) and hypermedia learning. Studies in epistemological beliefs, what individuals believe about the nature of knowledge (Hofer & Pintrich, 1997; Kardash & Scholes, 1996; Schommer, 1994, 1998), provide one lens through which the teaching-learning process in teacher education programs may be viewed. This study will report the results of a study on how teacher education students with different epistemological beliefs learn in a case-based hypermedia learning environment (CBHLE).

Personal epistemological development is a growing area of interest for psychologists and educators. These beliefs have been shown to influence students' learning in an ill-structured domain, problem-solving approach, and persistence when facing difficult tasks. With a synthesis of previous research, Schommer (1990) began to investigate epistemological beliefs in the late 1980s and later verified four dimensions of epistemological beliefs, including: (a) Ability to Learn; (b) Structure of Knowledge; (c) Speed of Learning; and (d) Stability of Knowledge. Other researchers have investigated the connection between the epistemological beliefs and the effects of these beliefs of learning in a hypertext system (Jacobson, Maouri, Mishra, & Kolar, 1996; Jacobson & Spiro, 1995; MacNeal, 2000). Jacobson and Spiro (1995) found that students who hold a more complex set of epistemic beliefs, or epistemological beliefs, were able to learn and apply their knowledge more efficiently after using the hypertext system than those who had only a simple set of epistemic beliefs. Other studies also found that students with less sophisticated beliefs were not likely to persist in information-seeking beyond the necessity (Tolhurst & Debus, 2002) and spent less time in completion of the hypertext activity (MacNeal, 2000). It is apparent that epistemological beliefs are not only important to general learning settings, but also to learning with computer-based materials. Therefore, additional empirical studies may be needed to substantiate the relationship between epistemological beliefs and learning in a hypertext or a hypermedia system.

A series of case-based hypermedia materials, *The Teacher Problem Solving Skills* (TPSS), has been produced by the Virtual Resource Center in Behavior Disorder (VRCBD), in order to enhance problem-solving skills of teacher education students preparing to serve children with emotional and behavioral disorders (Fitzgerald & Semrau, 1993-1997; 1998-2000). Through multimedia case materials, learners are able to learn complex case information, and practice skills needed in realistic situations.

In a recent review of hypermedia, Dillon and Gabbard (1998) examined the published findings from experimental studies of hypermedia learning outcomes. They suggested that certain individual characteristics play a role in the degree of success that a learner may achieve in hypermedia learning environments. However, learners' epistemological beliefs are one attribute nearly unexplored in empirical studies. There has been no study exploring the relationship between students' epistemological beliefs and their performance in a hypermedia or a case-based hypermedia learning environments. Previous studies on TPSS programs have shown that these programs offer an equally effective learning environment for many diverse learners, with the only effective difference among the learners' achievement being the total time engaged in the program. In other words, the more time learners spend in the hypermedia program, the better their knowledge acquisition and application in case artifacts. For that reason, a study is needed to evaluate the effects of learner epistemological beliefs, usage time, and learning outcomes with the TPSS materials. The research results may add substantiated empirical data to help teacher educators understand how their students learn from cases and hypermedia and provide information on effective implementation of case-based and hypermedia instruction. The following research questions were examined in this study.

1. Is there a relationship between teacher education students' epistemological beliefs and their usage of a case-based hypermedia learning environment in regard to total engagement time of the program?
2. Is there a relationship between teacher education students' epistemological beliefs and their learning outcomes in a case-based hypermedia learning environment in regard to the quality of case artifacts?

Methodology

The subjects in this study are sixty teacher education students from four universities enrolled in a special education course. Students used the TPSS program, *Instruction and Management in Emotional and Behavioral Disorders*, as an integrated course requirement. The hypermedia program was integrated into the16-week semester course. Students used the case "Amy" with background information, audio, videos, and a knowledge database. Through a variety of problem-solving activities, the participants take the role of Amy's teacher for planning instruction and creating cognitive-behavioral interventions to be used in her classroom setting. Problem-solving activities include identifying Amy's needs, assessing Amy's learning settings, designing management plans, planning therapeutic instruction, generalizing these plans into home and community, and evaluating these interventions according to best-practice prompts. The best-practice prompts are the guidelines normally used to evaluate the quality of intervention plans in the field of behavioral management. The participants completed the above activities as course assignments in their methods course. The participants' responses on the activities were stored on a floppy disk and submitted to the course instructors for later analysis. Figure 1 shows the "Case Manager" screenshot captured from this case-based hypermedia program.

Figure 1. Case Manager Screen of Amy Program

The instruments used in this study included: (a) The Epistemological Questionnaire (EQ) (Schommer, 1998); and (b) audit trail data collected within the hypermedia program. The study employed a combination of quantitative and qualitative data collection techniques including: (a) quantitative usage data collected via audit trail; (b) quantitative EQ scores; (c) quality analyses of the four problem solving activities (case consults, setting analysis, management plan and reflections, and instructional plan and reflections).

Independent Variables

The independent variables include four factor scores from the Epistemological Questionnaire.

Scores are used to represent the participants' epistemological beliefs in four areas: (a) Ability to Learn, (b) Structure of Knowledge, (c) Speed of Learning, and (d) Stability of Knowledge. The results of the EQ scores indicated that the subjects of this study had sophisticated views on three of the four epistemological belief factors.

Total Time Engaged in the CBHLE.

As the participants used the hypermedia system, an audit trail embedded in the program recorded the usage time and the navigational path. When the participants exited the program, the data were automatically collected and saved onto each participant's floppy disk. This information was then used to compile the total time spent in the program.

Case Artifacts

Four case artifacts were completed within the hypermedia program as part of the problem-solving activities for Amy. (a) Needs Analysis: The participant identified Amy's problems using a pre-formatted behavioral consulting process by considering her needs in ten areas, including academic, behavioral, affective, social skills, problem-solving skills, daily living, recreation, transition, home living and other areas; (b) Setting Analysis: The participant evaluated setting factors in Amy's classrooms and provided justification notes for each factor; (c) Management Plan and Reflection of the Plan: After defining Amy's needs, the participant developed management plans for Amy. Plans were entered into a template form and saved for later retrieval. The participant reflected on his or her prior management plans and provided responses to the four best practice prompts for management interventions; and (d) Instructional Plan and Reflection of the Plan: Instructional plans were entered into a template form and saved for later retrieval. The participant reflected on his and her prior instruction plans and provided responses to the four best practice prompts for instruction.

The analyses of data included frequencies and percentages, t-tests, Pearson's correlations and stepwise regression. Items on the EQ were compiled into 4 factors (Ability to Learn, Structure of Knowledge, Speed of Learning and Stability of Knowledge) in order to represent the epistemological beliefs of teacher education students. A set of correlation matrixes was used to analyze the relationships among variables before the regression technique was performed.

Results

The results showed a modest negative relationship (-.26 at .05 level of significance) between the belief in Speed of Learning and the total time engaged in the case-based hypermedia program. In other word, the more teacher education students believe in learning is quick, the less time they spend in the program. In addition, there is a moderate negative relationship among beliefs in Ability to Learn and scores in the case artifacts of Setting Analysis (-.31 at .05 level of significance) and Management Plan (-.38 at .01 level of significance). Conversely, there is a modest positive relationship between beliefs in Speed of Learning and Instructional Plan (.33 at the .05 level of significance). The regression analysis indicated that only 7% of the variation in Setting Analysis scores could be predicted from the belief in Ability to Learn. Only 12% of the variation in Management Plan scores could be predicted from the belief in Ability to Learn. The regression analysis showed that only 9% of the variation in Instructional Plan scores could be predicted from the belief in Speed of Learning.

Discussions

Based upon the above results, the more teacher education students believe that ability to learn is fixed, the poorer they perform in the problem-solving activities of Setting Analysis and Management Plan. The more they believe that when learning is quick, the better they perform in planning instructional intervention for the child featured in the case.

This study did not support earlier findings that belief in Speed of Learning might affect students' information processing and time spent in an ill-structured domain (Schommer, 1990; Jacobson et al., 1996; McNeal, 2000). This investigation revealed a weak relationship between belief in Speed of Learning and total time spent in the case-based hypermedia program. As noted by Schommer (1990), when learners encounter complex information, belief in quick, all-or-none learning appears to affect the extent to which they integrate knowledge. In other words, naïve learners in Speed of Learning are more likely to perform surface-processing strategies because they tend to make effortless rather than integrated solutions, and, consequently, spend less time to complete the required tasks. For the quick learner in a hypertext system, learning may imply acquiring information as quickly as possible in order to complete the task (MacNeal, 2000). However, the current study does not extend previous results gained by other researchers using paper-based materials to case-based hypermedia systems. More research is needed to address these inconsistent findings between these different delivery methods.

Data from the scores on the case artifacts indicated that belief in Ability to Learn had a negative relationship with the quality of work on some of the problem-solving activities embedded in this CBHLE. The more that teacher education students agreed that ability to learn is fixed at birth (higher score in Ability to Learn), the poorer the quality of their analysis on setting factors and their recommendations for managing Amy's behaviors from multiple perspectives to meet Amy's needs. Limited investigations have been conducted focusing on the relationships between belief in Ability to Learn and other aspects of learning. However, the findings of this study seem to support Schommer's study (Schommer, 1990) that learners with more sophisticated views toward Ability to Learn (lower score in Ability to Learn) tend to value persistence in hard work and in the face of academic difficulties; hence, they would be expected to achieve a better performance in problem-solving activities. Students with more sophisticated views toward Ability to Learn created more in-depth artifacts in Setting Analysis and generated more plausible behavioral management procedures to assist Amy.

The relationship between belief in Speed of Learning and the quality of work on intervention planning revealed inconsistent results from earlier research. Earlier studies (Schoenfeld, 1983; Schommer, 1990; 2002) suggest that naïve learners in Speed of Learning might perform more poorly because they are less inclined to engage and persist in difficult problems. Schraw, Dunkel, and Bendixen (1995) found that belief in Speed of Learning relates to students' problem solving in ill-defined content. However, this study did not replicate the afore-mentioned results for naïve learners in the problem-solving activity of instructional planning. In general, the structure and complexity of the Instructional Plan activity is parallel to that of Management Plan activity. Nevertheless, there are major differences in the content of the program that may account for this unexpected finding. Procedural information on behavior management is only provided through text in this CBHLE, while information on instructional planning is provided through video demonstration and expert commentary. The program features audio overviews of four curricular approaches (social skills training, conflict resolution, cognitive restructuring and anger control training) and classroom videos that demonstrate implementation of these approaches in the classroom. While it is true that the text information on behavioral management includes performance support tools and strategy information, these components are less accessible to the user. To utilize these resources, users must jump out from the main program; thus they are not easily available when users are constructing behavioral management artifacts. These two factors may contribute to difficulties of naïve learners in utilizing program resources to create good management plans. Even though naïve learners spent less time integrating ideas, they demonstrated better ability to utilize information in creating goals. This would be consistent in the cognitive load theory that suggests learning is enhanced when processing involves two or more channels in learning (Paivio, 1979, 1986).

Recommendations for Further Study

To address generalizability of the findings, it is recommended that future studies use a larger sample size and include a range of instructors who may not be experienced in teaching with hypermedia cases. Research needs to look at generalization across age and experience groups of students beyond those included in this study. Studies need to include participants representing the full spectrum of epistemological beliefs.

Previous research suggests that females are less likely to believe in Fixed Ability or Quick Learning (Schommer, 1993). Furthermore, the more students have to education, the less likely they are to believe in Fixed Ability or Quick Learning (Schommer, 1992). The current study found less than 10% of variation in learning outcomes could be explained by the self-reported epistemological beliefs. It is recommended that future studies might benefit from exploring demographic variables (i.e., gender and class rank) combined with epistemological beliefs on the usage and learning in case-based hypermedia programs. Such data would provide additional information for predicting the usage and learning outcomes of the program.

It is still unclear how well any paper-based instrument, such as the EQ or other available instruments, measures epistemological beliefs (Schommer, 2002; Schraw et al., 2002; Wood, Kitchener, Jensen, 2002). Future research should be conducted using multiple measurement approaches, such as unobtrusive audit trail analysis or in-depth interviews to accurately represent the complexity of epistemological beliefs. For example, some subjects revisited the program components of case information and electronic notes a greater extent than others; such differences in usage might reveal differences in reflecting or integrating information. Empirical studies suggest that learners with naïve views of epistemological beliefs tend to avoid integration during the problem-solving process (Schraw et al, 1995). In addition, conducting in-depth interviews or case studies on subjects' perception of textual and non-textual information could help to interpret unexpected performances in the case-based hypermedia program. Therefore, it appears that personal epistemological beliefs are so complex that mixed methodologies might provide better insight into the questions raised in this study.

Implications for Researchers and Instructors

Prior research shows that students with naïve belief in Speed of Learning tend to overrate their ability, knowledge and experiences (MacNeal, 2000; Schommer, 1990). Therefore, multiple approaches (observation, interview, and obtrusive audit trail) in combination with self-report instruments should be used to assess epistemological beliefs and to systematically explore the effects of these belief structures on knowledge acquisition and application in computer-based or non-computer-based learning systems (Jacobson et al. 1996).

Based on this study, it is apparent that students' epistemological beliefs have some effects on how learners apply knowledge in a case-based hypermedia program. When encountering ill-structured problems, students with naïve epistemological beliefs seem to have difficulties integrating ideas and perform poorly in problem-solving activities (Duell & Schommer-Aikins, 2001). Instructors might use knowledge of students' epistemological beliefs in designing instructional supports for students. For those students with more naïve beliefs, instructors might provide additional support to assist them in completing problem-solving activities, such as brief orientations to embedded scaffolds within the hypermedia program (Teacher Tools and Tool Resources) and rubrics to guide self-evaluation of work.

In order to motivate students and to better assess learning and performance using hypermedia case-based programs, instructors should integrate the materials and the content of the program into their course instruction in a meaningful way. Integration can occur through class discussions of the material and by engaging students in learning groups to work on the problem-solving activities from multiple perspectives. The use of the program can be made a course requirement with integration into the grading structure. Through these means, motivation can be improved externally through assignment of grades, collaboratively through class discussions for social construction of knowledge, and internally through personal reflection of performance and growth.

Acknowledgement

Appreciation is given to the four university instructors, though un-named to protect confidentiality, for their time and expertise and for allowing the researchers to probe their thoughts and learning outcomes with their students.

References

Dillon, A., & Gabbard, R. (1998). Hypermedia as an educational technology: A review of the quantitative research literature on learner comprehension, control and style. *Review of Educational Research, 68*(3), 322-349.

Duell, O. K., & Schommer-Aikins, M. (2001). Measures of people's beliefs about knowledge and learning. *Educational Psychology Review, 13*, 419-449.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(Spring), 88-140.

Fitzgerald, G., & Semrau, L. (1995-1998). *Enhanced Training Program in Behavior Disorders: Skills for Inclusion Settings.* United States Department of Education: Office of Special Education Program.

Fitzgerald, G., & Semrau, L. (1998-2000). *Virtual Resource Center in Behavioral Disorders* (Project #H029K70089). United States Department of Education: Office of Special Education Program.

Jacobson, M. J., Maouri, C., Mishra, P., & Kolar, C. (1996). Learning with hypertext learning environments: Theory, design and research. *Journal of Educational Multimedia and Hypermedia, 5*(3/4), 239-281.

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Multimedia and Hypermedia, 12*(5), 301-333.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology, 88*(2), 260-271.

MacNeal, L. G. (2000). *Epistemic beliefs and comprehension in a hypertext system.* Widener University, Chester, PA.

Paivio, A. (1979). *Imagery and verbal processes*. Hillsdale, NJ: Lawrence Erlbaum.

Paivio, A. (1986). *Mental representation: A dual coding approach*. New York: Oxford University Press.

Schoenfeld, A. H. (1983). Beyond the purely cognitive: Belief systems, social cognition, and metacognitions as driving forces in intellectual performances. *Cognitive Science, 7,* 239-363.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(Sept), 498-504.

Schommer, M. (1992). Epistemological Beliefs and Mathematical Text Comprehension: Believing It Is Simple Does Not Make It So. *Journal of Educational Psychology, 84(4)*, 435-443.

Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*(3), 406-411.

Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P. A. Alexander (Eds.), *Beliefs about Text and Instruction with Text* (pp. 25-40). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology, 68*, 551-562.

Schommer-Aikins, M (2002). An evolving theoretical framework for an epistemological belief system. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (Vol. 1, pp. 103-118). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9*, 523-528.

Tolhurst, D., & Debus, R. L. (2002). The influence of prior knowledge, attitude, ability and activity structure on students' learning and use of software. *Journal of Educational Computing Research, 26* (in press).

Wood, P., Kitchener, K., & Jensen, L. (2002). Considerations in the design and evaluation of a paper-and-pencil measure of epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (Vol. 1, pp. 277-294). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Understanding Teachers' Use of the Web to Support Higher Levels of Thinking: Exploring the Relationship of Internet Teaching Efficacy and Purposes

David Pratt
Purdue University

Abstract

This study examined teachers' self-efficacy and purposes for using the World Wide Web with students. Participants for the study included twenty-two teachers who were found to use the Internet as a tool to support student learning but differed in a specific form of self-efficacy known as Internet Teaching Efficacy (ITE). Interview and survey data analysis revealed a significant correlation between teachers' ITE and their educational purposes. Specifically, teachers with less confidence in using the Internet (lower ITE) described the use of the World Wide Web as supporting pupil participation and lower level cognitive thought. Teachers with higher ITE reported a much larger percentage of purpose statements supporting independent, creative and self-regulated learning at higher cognitive levels. Implications for the preparation of teachers are discussed in light of these results.

While the problems with student access to computers in schools have decreased, research has provided evidence that teachers' use of the computer as a tool to support higher levels of student learning has yet to be realized (Becker, 1999). Recent studies have demonstrated that teachers assign students to use the computer more for drill and practice activities than for research, inquiry or group activities (Maddux, 1998; Williams, 2000). Cuban (2001) pointed out that regardless of the abundant availability of computers, software, and professional development in schools, traditional teaching methodologies have not been altered.

Researchers continue to suggest, however, that access to the resource of the World Wide Web has the potential to promote higher levels of thinking among students (Doherty, 1998; Hopson, et. al, 2002), and is a tool conducive for inquiry-based learning (Love & McVey, 2000). The importance of considering teachers' purposes for using the Internet are therefore important to consider in light of this distinction made in the literature.

Review of Literature

Educational Purposes

Many researchers have acknowledged the importance of considering a teacher's sense of educational purposes as a component of their professional development (Gore & Zeichner, 1991; Grossman & Richert, 1988; Shulman, 1987). More recently, Copeland & D'Emidio-Caston (1998) found that a teacher's beliefs about educational purposes are critical for understanding teachers' practical theory. One's practical theory is thought of as a set of conceptualizations, theories, beliefs, and values that have developed in the course of learning to teach. It is assumed that teachers utilize this set of beliefs to make meaning out of classroom events and to make decisions based on that meaning.

For many years, emphasis has been placed on the importance of considering higher level purposes together with lower ones. Much of this focus has been connected with Bloom's taxonomy. For example, several studies have been conducted to suggest that teachers should use this taxonomy as a template for writing questions in order to ensure that diversity of questions and critical thinking skills are involved in lessons (Callison, 1998; Gilbert, 1992). Other studies specifically describe the importance for considering higher level thinking skills as a necessity for survival in a rapidly changing world (Lee & Dinkins, 1998; Paul, et. al., 1990). Understanding teachers' educational purposes for using the World Wide Web may provide a better understanding of why teachers' use technology in the classroom the way they do. Maddux (1998) includes in his study, a description of teachers using technology to emphasize whole class teaching and telling practices that focused on factual and rote learning. Based on these observations, Maddux argues that the underlying purposes these teachers had for use with students, limited the potential of the Internet as a tool to support higher levels of thinking.

Self-Efficacy

Consideration of teachers' self-efficacy may be helpful in trying to understand the quality of teachers' integration of Internet use with students. Bandura defined self-efficacy as "belief in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p.3). According to Bandura, it is not enough to be knowledgeable about teaching or even possess the skills needed to teach: A teacher must also believe in his or her own ability to perform the actions needed to accomplish a particular behavior.

Bandura also claimed that self-efficacy is task-specific. That is, one's confidence is related to a certain task under particular conditions. Personal Teaching Efficacy, a specific form of self-efficacy, has been correlated with many teacher behaviors including feedback patters, interactions with students, student achievement, and innovativeness in pedagogy (Ashton & Webb, 1986; Ghaith & Yaghi, 1997; Gibson & Dembo, 1984; Woolfolk & Hoy, 1990). The present study uses a task-specific scale to measure teachers' confidence in using the Internet with students. This is important to consider as a teacher's confidence for teaching in general may differ from his or her confidence when teaching using the Internet. Teachers' Internet Teaching Efficacy (ITE) is being assessed for the present study and defined as a belief in one's ability to use the Internet as a tool to help students learn.

Purposes of the Study

The present study proposes that in order to understand teachers' use of the World Wide Web with students, one must consider teachers' confidence and purposes for using it with students. Therefore, this study focuses on answering the following two research questions: 1) What type of educational purposes do teachers have for using the World Wide Web with students? 2) What is the nature of the association, if any, between teachers' confidence (ITE) and purposes for using the World Wide Web with students?

Data Sources and Evidence

A survey was given to 120 teachers to assess their efficacy for using the Internet with students. The Personal Internet Efficacy Beliefs Scale was developed by Koul & Rubba (1999) and was previously used and tested for its validity and reliability (Appendix A). This initial survey also collected data about each teacher's amount of Internet use with students. A criterion sample of twenty-two teachers who used the Web with students at least three times during the year and who varied in their ITE scores were selected as participants for the next phase of study. A semi-structured interview guide was then used to elicit information for three general categories: background information, context of use, and purposes for using the Internet with students.

The present study uses Copeland and D'Emidio-Caston's (1998) established categories for coding and analyzing of interview data. The studies conducted by Copeland and his colleagues (1994; 1998) reflect this notion of different purposes by type and level. Using a Constant Comparative Analysis approach they formed categories with the interview data that distinguished between expressions of purposes held by the teacher. They found these categories helpful in labeling statements by specific characteristics and for use in further analysis. There emerged two basic categories of purpose among the pupil-related statements:

- those related to the classroom participation behavior of pupils and to their lower cognitive thought (labeled by the author as narrow purposes).

- those related to the development of pupils as independent, creative and self-regulated learners (labeled by the author as broad purposes).

Under the two main categories of broad and narrow purposes there were found sub-categories that relate to specific types of statements made about students. These include Pupil Comportment, Pupil Participation, and Lower Level Cognitive thought under Narrow Purposes; Broad Purposes included higher-level cognitive thought, affective thought, and development of learner autonomy (Refer to Appendix B for a complete description of each sub-category for broad and narrow purposes).

Modification of Original Sub-Categories

This established coding procedure was useful in categorizing teachers' purposes for using the World Wide Web in their teaching, however, Respondents' descriptions in the present study elicited one additional purpose unrelated to the original categories developed by Copeland & D'emidio-Caston (1998). Many teachers included in their purpose statements the notion that students should only be allowed access to specific sites which led to information that the teacher determined as valuable. These expressions were labeled as a new sub-category "Following Directions". This sub-category was classified as "narrow" since it represented a teacher's concern for students to focus on specific teacher-controlled information (see Appendix B sub-category 3).

A total of twenty-two teachers were selected as Respondents for the present study representing five different grade levels and eight different schools. These teachers also represented a wide range of teaching experience (2-34 years). The total number of Respondents averaged 14.1 years of teaching. Table 1 presents these Respondents by school, grade level, years taught, and corresponding ITE scores. (Note: All teachers and school names are fictitious).

To assist in later analysis of Respondents, it will be helpful to distinguish between two groups: those teachers with "lower ITE" and those with "higher ITE". Therefore, the Respondents' position in Table 1 was determined by the rank order of their ITE scores. Two criteria were used to determine the dividing point between the two groups. First, Respondents were separated by scores above and below the mean (4.09). Twelve teachers reported scores below the mean, versus ten that were above. Secondly, a natural gap of scores presented itself between Cathy (4.31) and Colleen (4.00). This was one of the largest disparities between scores.

Table 1. *Summary of Respondents Selected for Study*

Name	School	Grade	Year	ITE Score
		Higher ITE Teachers		
Mary	Adams	5	25	5.85
Jenny	Hoover	4	18	5.69
Karen	Hoover	6	15	5.69
Lynne	Hoover	5	2	5.23
Kris	Adams	5	18	5.15
Haley	Madison	6	4	5.00
Alice	Harding Jr. High	7/8	6	4.54
Marta	Pine	6	6	4.46
Sammy	Hoover	5	15	4.38
Cathy	Hoover	6	19	4.31
		Lower ITE Teachers		
Colleen	Columbia Jr. High	7/8	7	4.00
Michelle	Adams	4	9	4.00
Monica	Monroe Jr. High	7/8	10	3.92
Chris	Adams	5	8	3.85
Elizabeth	Columbia Jr. High	7/8	6	3.69
Candy	Kennedy	4	6	3.38
Susan	Hoover	4	10	3.31
Caroline	Pine	6	34	3.23
Myra	Columbia Jr. High	7/8	12	2.85
Jamie	Madison	4	27	2.81
Donna	Columbia Jr. High	7/8	25	2.54
Summer	McManus Middle	6	20	2.31

Findings

Analysis of data from interviews of the twenty-two teachers revealed a total of 202 purpose statements related to student learning involving the Internet. Recall that these were categorized into two major types of statements: narrow and broad. The smallest number of statements for any teacher was six while the largest was 13. Table 2 provides the purpose statements made by each Respondent, classified as belonging to each sub-category and whether that category is broad or narrow. Each Respondent's Broadness Score was calculated by dividing the number of broad purposes by the Respondents' total number of purposes.

Table 2. *Respondents' ITE Scores, Number of Purpose Statements by Type, and Resulting Broadness Scores*

Respondent	ITE	Purpose Statements									Broadness
		Narrow					Broad				
Name	Score	1	2	3	4	Total	5	6	7	Total	Score
						Higher ITE Teachers					
Mary	5.85	0	0	0	3	3	2	1	0	3	.50
Jenny	5.69	0	0	0	4	4	2	1	0	3	.43
Karen	5.69	0	2	2	2	6	2	3	2	7	.46
Lynne	5.23	0	0	0	3	3	1	4	1	6	.67
Kris	5.15	0	1	0	2	3	4	2	0	6	.67
Haley	5.00	0	1	0	4	5	1	3	0	4	.44
Alice	4.54	1	0	1	3	5	3	0	1	4	.44
Marta	4.46	0	2	0	2	4	2	1	1	4	.44
Sammy	4.38	0	0	0	3	3	2	3	2	7	.70
Cathy	4.31	0	0	0	3	3	2	4	1	7	.70
						Lower ITE Teachers					
Colleen	4.00	0	0	0	6	6	1	2	1	4	.40
Michelle	4.00	0	1	0	3	4	0	2	0	2	.33
Monica	3.92	0	1	3	1	5	3	1	0	4	.40
Chris	3.85	0	0	1	7	8	0	2	0	2	.20
Elizabeth	3.69	0	1	0	7	8	2	1	0	3	.30
Candy	3.38	1	0	1	4	6	2	0	1	3	.33
Susan	3.31	0	0	1	6	7	0	1	0	1	.13
Caroline	3.23	0	2	1	5	8	2	0	0	2	.20
Myra	2.85	1	0	1	3	5	1	0	0	1	.17
Jamie	2.81	0	3	0	5	8	0	3	0	3	.27
Donna	2.54	0	0	1	7	8	1	1	0	2	.20
Summer	2.31	0	2	1	5	8	1	1	0	2	.20

1 = Pupil comportment; 2 = Pupil Participation 3 = Following Directions 4 = Lower Cognitive Thinking
5 = Higher Level Thinking 6 = Affective Thought 7 = Autonomous Learning

The data depicted in Table 2 were submitted to a Spearman Rho correlation test using SPSS software to determine the type of relationship that exists between the two sets of scores (ITE scores and Broadness Scores). Analysis of data revealed a positive correlation significant at .01 (r=.817). Figure 1 provides a scatterplot of each Respondents scores while labeling each of them as belonging to either the higher or lower ITE group.

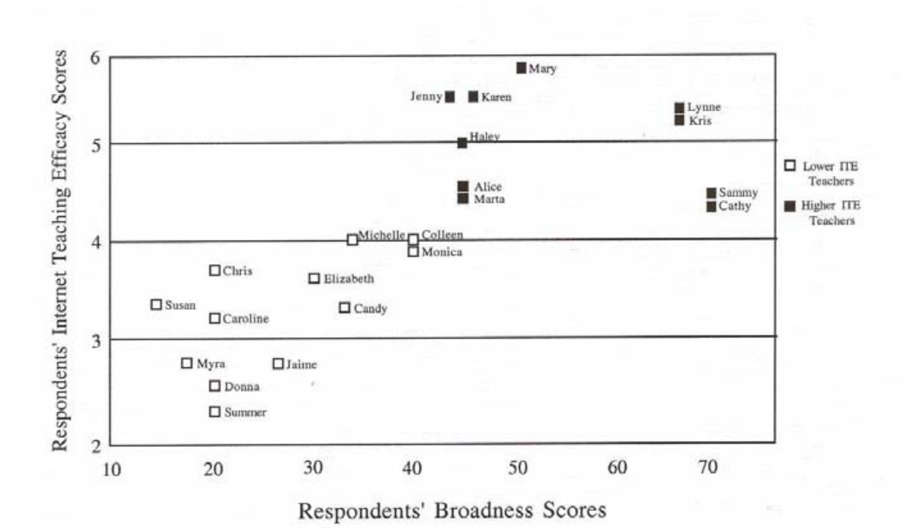


Figure 1. Scatter plot of Participants' ITE and Broadness Scores

While a significant correlation between ITE and purposes is important to recognize, one can use interview data to confirm this result and to provide a more specific examination of the nature of such an association. In order to accomplish this, higher ITE teachers were compared to lower ITE teachers with regards to their specific purposes. Figure 2 provides an analysis of the differences between higher and lower ITE teachers for each of the sub-categories.

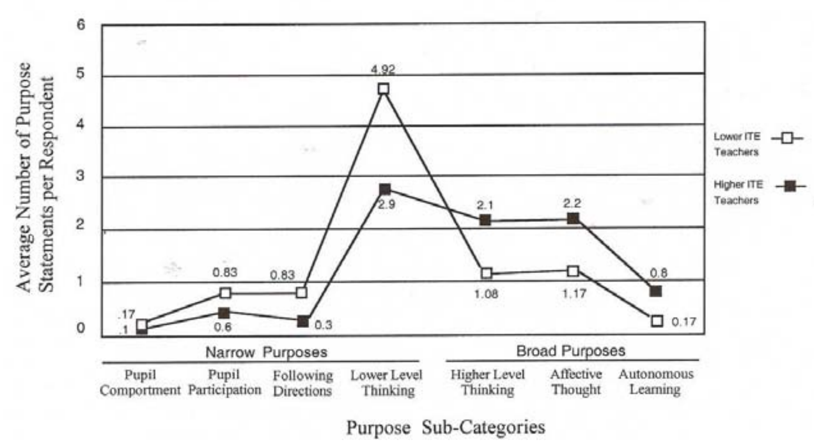


Figure 2. Average Number of Purpose Statements per Respondent by Sub-Category

Notice that in each of the narrow purpose sub-categories, there was a greater average of statements provided by teachers with lower ITE than the higher ITE group. A concern for lower level cognitive thought was the most common sub-category mentioned and teachers' responses were found to be substantially different between groups as there were nearly twice as many statements from the lower ITE group than the higher one. While lower ITE teachers averaged nearly five statements regarding lower level thinking (4.92), higher ITE teachers included an average of nearly three (2.90).

In contrast to that, higher level thinking purposes were mentioned nearly twice as many times by higher ITE teachers than by lower ITE teachers (2.10 vs. 1.08). Higher ITE teachers were also found to mention an average of one more statement regarding Affective Concerns when compared to the lower ITE teachers (2.20 vs. 1.17). Differences were also found between groups in the "Following Directions" and "Autonomous Learning" sub-categories. Less significant differences were found between groups in sub-categories "Pupil Comportment" and "Pupil Participation".

Importance of the Study

At the center of this study was an attempt to understand why teachers use the Internet with students and what contributes to those decisions. The findings of the present study suggest an association between one's Internet Teaching Efficacy (ITE) and one's purposes for using the Internet with students. Specifically, teachers with higher ITE included more statements of purposes categorized as 'broad' (focused on higher level thinking, etc.) while those with lower ITE tended to include a greater proportion of 'narrow' purposes when describing their use of the web with students. Understanding the relationship that exists between teachers' Internet Teaching Efficacy and their purposes for using the Internet with their students may be useful for professional development and teacher education programs when considering how to best support teachers in using the Internet for broader purposes. More generally, this issue is important to consider in terms of the potential of schools' investment related to technology and its potential for helping prepare students to become prepared for the Knowledge Age.

Teacher education programs may need to consider that providing a strong knowledge base of content and pedagogy may not be sufficient to help teachers grow and succeed as professionals. Several researchers have pointed to the importance of explicitly acknowledging those beliefs at this level (Ashton & Webb, 1986; Hollingsworth, 1986; Ross, 1995). Additionally, professional development leaders may need to consider specifically addressing teachers' self-efficacy during inservices. If the purpose for teachers' use of technology depends not just on effectiveness of skill but of efficacy, then one would assume that, along with workshops focused on skill proficiency, support for specifically increasing one's computer or Internet efficacy may be important. Clearly, there are limitations to this exploratory study and more work needs to be done in this area. First, understanding the sources contributing to the teachers' Internet Teaching Efficacy may be helpful, especially when considering its practical application to teacher preparation. For example, if one knew that vicarious experience (modeling) is the most common way teachers increase their self-efficacy, one could focus on providing good examples of technology use in classrooms in order to increase teachers' self-efficacy. In addition, it would be helpful for professional development and preservice programs to be aware of conditions that fostered higher levels of efficacy for using the Internet. While the present study was careful to assess teachers' efficacy based on similar access and support for technology use, understanding specific factors that influence Internet Teaching Efficacy would be valuable.

Finally, considering the cost and potential of having Internet-compatible computers suggests that their use should be greater than use for drill and practice. Billions of dollars have been spent in recent years to assure that schools are connected to the vast resource of the World Wide Web (Ballard, 2000). Considering the cost involved with such a resource, many agree that teachers should employ methods to increase the use of technology as a tool to support higher order thinking skills with their students. This study supports the notion that teachers can include broader purposes among the narrower ones when using the Internet as a tool to help students learn.

The present study suggests that using the framework of self-efficacy theory we may be able to better understand the purposes teachers have for using the Internet in the classroom. The study also demonstrates that teachers are able to use the Internet to incorporate broader purposes among others that they may have for a lesson. Further research following this line of inquiry may help teacher preparation programs, inservice developers, policy makers, and administrators become more prepared to make better decisions in guiding teachers to improve the quality of their use.

References

Ashton, P.T., & Webb, R.B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.

Bandura, A. (1997). *Self-efficacy : The exercise of control*. New York: W.H. Freeman.

Becker, H. J. (1999). Internet use by teachers: Conditions of professional use and teacher-directed student use. Teaching, Learning, and Computing: 1998 National Survey Report #1. Center for Research on Information Technology and Organizations. [Online] Available: http://www.crito.uci.edu/TLC/findings/Internet-Use/startpage.htm.

Callison, D. & Tilley, C. (1998). Information and media literacies: Towards a common core. *School Library Media Activities Monthly*, 15 (2), 25-28.

Copeland, W.D., Birmingham, C., DeMeulle, L., D'Emidio-Caston, M., Natal, D. (1994). Making meaning in classrooms: An investigation of cognitive processes in aspiring teachers, experienced teachers, and their peers. *American Educational Research Journal*, *31*, 167-195.

Copeland, W. D., D'Emidio-Caston, M. (1998). Indicators of development of practical theory in pre-service teacher education students. *Teaching and Teacher Education*, *14*, 513-534.

Cuban, L. (2001). *Oversold and Underused: Computers in the classroom.* Cambridge: Harvard University Press.

Doherty, A. (1998). The Internet: Destined to become a passive surfing technology? *Educational Technology*, September-October, 61- 63.

Ghaith, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 13*, 451-458.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation*. Journal of Educational Psychology, 76*, 569-582..

Gilbert, S.W. (1992). Systematic questioning: Taxonomies that develop critical thinking skills. *Science Teacher*, *59* (9), 41-16.

Gore, J. M. & Zeichner, K. M. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. *Teaching and Teacher Education*, *7* (2), 119-136.

Grossman, P. L. & Richert, A. E. (1988). Unacknowledged knowledge growth: A re-examination of the effects of teacher education. *Teaching and Teacher Education: An International Journal of Research and Studies, 4* (1), 53-62.

Hopson M. H., Simms, R. L., & Knezek, G. A. (2002). Using a technology-enriched environment to improve higher-order thinking skills. *Journal of Research on Technology in Education, 34* (2),109-119.

Koul, R. and Rubba, P. (1999). An analysis of the reliability and validity of personal internet teaching efficacy beliefs scale. *Electronic journal of science education*, [Online] 4. Available: http://unr.edu/homepage/crowther/ejse/koulrubba.html.

Lee A. A. (1998). *Critical thinking for the new millennium: A pedagogical imperative.* ERIC No. ED 429 913.

Love, R. & McVey, M. (2000). Teachers' use of the Internet. *Teachers college record*. ID 10538. [Online]. Available: http://www.tcrecord.org.

Maddux, C.D. (1998). Barriers to the successful use of information technology in education. *Computers in the Schools*, 14, 5-11.

Paul, R. W., Binker, A. J. A. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world.* ERIC No. 338 557.

QED Report (2000). Quality Education Data, Inc. as reported in USA Today, 2/28/2000 by Michelle Healy [Online]. Available: htttp://www.infotoday.com/MMSchools/nov00/ballard.htm#refs

Ross, J. A. (1995). Strategies for enhancing teachers' beliefs in their effectiveness: Research on a school improvement hypothesis. *Teachers College Record*, *97*, 227-251.

Shulman, L. S. (1987). Assessment for teaching: An initiative for the profession. *Phi Delta Kappan*, *69*, 38-44.

Williams, C. (2000). *Internet access in U.S. public schools and classrooms: 1994-1999 (NCES 2000-086).* U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Woolfolk, A. E., Rosoff, B., & Hoy, W.K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education, 6*, 137-148.

Appendix A. Personal Internet Teaching Efficacy Beliefs Scale

Answer on a scale from Strongly Disagree to Strongly Agree.	**Strongly disagree**	**Moderately disagree**	**Disagree slightly more than agree**	**Agree slightly more than disagree**	**Moderately agree**	**Strongly agree**
I am continually finding better ways to teach with the Internet.	O	O	O	O	O	O
Even when I try very hard, I do not teach as well using the Internet as I teach in other ways.	O	O	O	O	O	O
I know how to teach effectively using the Internet	O	O	O	O	O	O
I am not very effective in monitoring activities that involve using the Internet	O	O	O	O	O	O
I generally teach ineffectively when using the Internet.	O	O	O	O	O	O
I understand how to use the Internet well enough to be effective in teaching with it.	O	O	O	O	O	O
I find it difficult to explain to students how the Internet works.	O	O	O	O	O	O
I am typically unable to answer students' Internet questions.	O	O	O	O	O	O
I wonder if I have the necessary skill to teach using the Internet.	O	O	O	O	O	O
Given a choice, I would not invite the principal to evaluate my teaching when I use the Internet in a lesson.	O	O	O	O	O	O
When teaching using the Internet, I usually welcome student questions.	O	O	O	O	O	O
I don’t know what to do to turn students on to using the Internet.	O	O	O	O	O	O
When a student has difficulty understanding how to use the Internet, I am usually at a loss as to how to help the student understand it better.	O	O	O	O	O	O

Appendix B. Categories of Purpose Statements

I. Pupil-Oriented Purposes

Narrower Purposes - Those Related to Pupil Behavior and Lower Cognitive Thought

Sub-Category 1. Pupil Comportment
Concern for how students are "behaving themselves" in the classroom. Includes such things as off-task behavior, paying attention, fidgeting, moving about without permission and sitting quietly.

Sub-Category 2. Pupil Participation in the Lesson
Concern for the manner and patterns of student involvement in the publicly experienced classroom conversation. Includes patterns of obtaining opportunities to speak, factors that encourage or inhibit involvement and patterns of involvement which are apparently governed by factors such as gender, race or physical location.

**Sub-Category 3. Following Directions*
Concern for how well students can follow specific directions given for a certain task. Includes such things as not letting students search on their own or go off on uncharted areas of the web. Their concern is that students will get lost or access inappropriate material.

Sub-Category4. Lower-level Cognitive Thought
Concern with students' cognitive thought typified by simple recall of information, work with information that is readily at hand, and the changing of the form of information without changing its meaning (the lowest two levels of Bloom's Cognitive Taxonomy). Includes such thoughtful actions as locating specific information in a book, learning specific information, and reading from and comprehending text.

Broader Purposes - Those related to the Development of Independent, Creative and Self-Regulated Learners

Sub-Category 5. Higher-Level Cognitive Thought
Concern with students' cognitive thought typified by more complex manipulation and use of information (the higher levels of Bloom's Cognitive Taxonomy). Includes determining whether a particular piece of information supports a particular statement, analysis of information into its component parts, creation of new ideas such as hypotheses, and evaluation of the truth or falsity of ideas.

Sub-Category 6. Affective Thought
Concern with students' thought related to attitudes and values, whether found currently in students or the purpose of educational activities. Affective thought may be expressed in relation to either of two types of concerns.

Pupil Affective Dispositions. A concern with the agency of the individual student, i.e., is he or she willing to engage in the learning activity? This is a volitional concern at the individual level. Includes such behaviors as enjoying or taking pride in participation and being excited about learning.

Conditions that Reflect or Even Promote Affective Thought. A concern with the affective condition of a student or students. This may effect the larger environment in which students work and its promotion of affective thought. Includes the teacher's assumption of a supportive attitude, a comfortable rapport between the teacher and students and students feeling reassured that participation will be accepted.

Sub-Category 7. Development of Learner Autonomy.
A concern with the development of student's ability and inclination to proceed independently in a learning task. Includes appropriate self-concept (self-esteem), self-monitoring skills and a feeling of efficacy (self-empowerment).

Other Pupil -Oriented Purposes

Sub-Category 8. Other
Student behaviors that do not fit into the above six categories.

II. Teacher-Oriented Purposes

Sub-Category 9. Teacher Behavior
Any reference that focuses on the behavior of the teacher.

Categories of Purpose Statements: from Copeland & D'emidio-Caston (1998)
* category added based on present study's interview data analysis

When Do Headings Improve Learning? A Synthesis of Verbal Signals Research

Gustavo E. Prestera
effect*Performance*, Inc.

Roy Clariana
Penn State University

Abstract

Headings are commonly-used verbal signals that divide instructional text into smaller sections and prime *the reader to the main ideas discussed in those sections. Writers of instructional text may incorporate headings into their work without realizing that it impacts the reader's ability to remember and process the information. In this review of literature, the effects of headings and other verbal signals on learning are explored. Of the studies reviewed, 15 are discussed and synthesized to form practical guidelines for instructional designers and educators. Among the findings, headings have been shown to promote recall of main ideas, knowledge of organizational structure, and problem-solving. However, even as signaling makes key topics more memorable, it can also inhibit recall and comprehension of subordinate information that is not signaled. These and other findings are presented and discussed.*

Introduction

Headings represent one member of a large family of instructional variables called *signals*. Verbal signals include titles, introductions, overviews, headings, subheadings, transition statements, and summaries. Non-verbal typographical signals include bolding, italics, underlining, spacing, all-caps, font styles, font sizes, color text, and color highlighting. In instructional settings, verbal and non-verbal signals help learners navigate bodies of instructional text by providing structure, sequence, organization, and flow. Signals are used in the expectation that they make the relevant topics stand out in importance and thereby more memorable (Mautone & Mayer, 2001). Headings, like other verbal signals, are "non-content words (that) make the semantic content and structure of expository paragraphs more explicit" (Mautone & Mayer, 2001: p. 377; See also Meyer, 1975). Headings divide the instructional text into smaller chunks, presumably making the information easier to encode. Readers can make use of headings as priming aids *before* reading (to help them skim the passage), as encoding aids *during* reading (to orient themselves to the text that follows), as retrieval aids *after* reading (to cue retrieval of a particular concept), and as search aids (to assist in locating specific information). Signals research has typically focused on the *encoding* effects of headings (Hartley & Trueman, 1985).

The encoding effects of headings have been shown to vary by a number of factors (Hartley & Trueman, 1985; Marcinkiewicz & Clariana, 1997). For example, Meyer, Brandt, and Bluth (1980) found that headings in text enhanced recall for poor readers but not for good readers. Conversely, Hartley and Trueman (1985) found that heading statements aided high-ability students most and low-ability students least. Such differences are common in headings research and are, in part, the result of differences in the treatment and assessment materials. However, these inconsistencies are also indicative of the subtle yet complex relationship that exists between the internal memory processes of encoding and retrieval and the perception of signals. The following is a review of the most pertinent studies.

Review of Literature

In the absence of headings and other signaling devices, which highlight the intended organizational structure of the reading passages, the reader utilizes a linear approach to reading the text. The ideas are stored in memory in the form of a serial list. The learner's processing of instructional content and his/her ability to recall or summarize key concepts is then highly dependent upon the content's serial positioning (Loman & Mayer, 1983), the textual elaboration of key points (Lorch, Lorch, & Inman, 1993), and the learner's prior knowledge of the subject matter (Lorch & Lorch, 1996b). These factors are "secondary indicators of importance of text topics" (Lorch, Lorch, Ritchey, McGovern, & Coleman, 2001:

p. 172). The presence of signals, however, can mitigate these factors, by focusing the learner's attention on the organizational structure of the content. Signaling devices enable the reader to use a structure strategy in which ideas are associated with higher-level topics within a broader organizational structure (Sanchez, Lorch, & Lorch, 2000). This should make it easier for readers to access ideas even when those ideas are not positioned at the beginning or end of the text. Signaling should improve recall of those ideas.

Despite this common sense hypothesis, Lorch and Lorch (1996b) note that some prior studies have failed to find significant signaling effects on total recall, i.e., the overall number of idea units recalled (Brooks, Dansereau, Spurlin, & Holley, 1983; Lorch et al., 1993; Meyer, 1975; Meyer & Rice, 1982). Even so, headings and other signaling devices have been shown to improve recall under certain conditions, for certain types of information, and for certain learners.

Do signaling effects vary by signal type?

Effects of headings, summaries, outlines, and underlined text on comprehension. In a study involving 432 Air Force basic trainees, Christensen and Stordahl (1955) examined the effects of signals on comprehension by presenting students with an outline, a summary, underlining, and/or paragraph headings. Using a factorial design, the researchers examined 35 different possible combinations of signals in comparison with a no signals control group ($n = 12$ per group). The pretest and posttests used consisted of multiple-choice test items. No significant interactions were uncovered for either pretest or posttest scores. Neither headings alone nor in combination with other signals had any appreciable impact on comprehension.

Effects of headings, previews, and connectors on topic recognition. In an effort to isolate the effects of headings, previews, and logical connectors on the retrieval of both inferential and detailed information, Spyrikakis and Standal (1987) examined all three signal types individually and in tandem with each other. The study involved 300 college students who each read one of four passages (topics were nitrates, corrosion, algae, and biomedical) of varying structure, depth, and length (ranging from 362 to 766 words). Each also completed one of four 10-item multiple choice tests measuring the ability to recognize both super-ordinate topics and subordinate information in the passages. The use of headings was found to be significantly beneficial with only one of the four passages (biomedical) and then only with regards to enhancing super-ordinate/inference comprehension, $F(1, 64) = 4.19$, $p < .05$ (p.291). The biomedical passage was the longest (600-766 words), the richest in terms of the number of topics (42 versus 36 for the next highest), and the most difficult in terms of grade level. This implies that headings are most useful for stimulating recognition of key points (super-ordinate topics) when the content is long, difficult, and complex. Nevertheless, Spyrikakis and Standal (1987) suggest that in this study, "the results are suspect due to a ceiling effect" (p. 291).

Effects of headings and outlines on free recall. Krug, George, Hannon, and Glover (1989) conducted a series of experiments testing the impact of presenting outlines (prefacing text) and headings (interspersed within the text) at encoding on students' free recall of prose. In the first experiment, involving 62 undergraduate Education majors and a 600-word essay containing 48 idea units, Krug et al. found that both headings and outlines improved recall equally well and that a combined headings/outline approach was the most effective, $F(3, 58) = 10.44$, $p < .01$. In their second experiment, Krug and his colleagues asked 56 students to read a 5500-word textbook chapter – in an effort to confirm the results of the first experiment in a more realistic classroom context – and again, headings and outlines improved recall equally well. The combined approach of headings with outlines was the most effective.

Effects of headings, overviews, and summaries on free recall. In a study designed in part to isolate the effects of different signaling devices, Lorch and Lorch (1995) tested the individual effects of headings, overviews, and summaries on free recall of main and subordinate topics. College students ($n = 274$) were given a 1,750-word passage that varied in terms of the presence or absence of the three signaling devices, with 31 of those participants receiving the headings passage. All three signaling devices improved recall of the 12 main topics (topic access), $F(1, 176) = 37.39$, $p < .05$, and conditional recall of subordinate topics, $F(1, 176) = 12.29$, $p < .05$ (p. 541). All three signals enhanced recall equally well. As a result, the results of all three signaling conditions were combined and analyzed with relation to the presence or absence of recall cues. Consistent with the encoding specificity hypothesis (Tulving & Thomson, 1973), Lorch and Lorch found that the presence of signals at encoding (in the passage) and recall cues (during the free recall task) significantly enhanced recall of both main topics and subordinate information. Interestingly, in the absence of recall cues, signaling only during encoding depressed recall scores, although these effects did not reach significance. The encoding specificity hypothesis predicts that context cues presented at encoding are most effective only when they are presented again at the time of retrieval. The findings in the Lorch and Lorch (1995) study suggest that signals operate in a similar manner as other context cues, enhancing recall the most when present at both encoding and retrieval.

Do signaling effects vary by ability level?
Signaling effects on cued recall by ability level. In a study measuring the impact of titles and paragraph headings on recall of prose text, Hartley, Kenely, Owen, and Trueman (1980) found that using titles and headings improved recall on 8 short-answer questions. The 175 second-graders who participated were blocked by ability level; asked to read a 400-word passage with (A) no title or headings, (B) with title but no headings, (C) with title and underlined passage headings written in the form of statements, or (D) with title and underlined passage headings written in the form of questions. Although they did not find significant recall differences between (C) headings worded as statements and (D) those worded as questions, Hartley et al. did find that conditions C and D combined – regardless of format – prompted significantly higher recall than conditions A and B combined, $F(3, 171) = 4.88$, $p < .01$. This suggests that headings, regardless of the format, improve the recall of prose. A significant interaction was also found between the use of headings and ability level. Low ability students benefited more from headings than high ability students, $F(6, 151) = 2.47$, $p < .05$ (p. 306).

Do signaling effects vary by item difficulty, familiarity, or prior knowledge?
Headings effects on retention and prior knowledge. Wilhite (1988) found a significant interaction between prior knowledge (1988) and retention. The study involved 116 college students who were asked to read an 8-section, 1,760-word passage with or without headings and answer 24 multiple-choice questions (8 low-level topic pretest items, 8 low-level topic posttest items, and 8 high-level posttest items). When means were adjusted with the prior knowledge covariate (pretest score), headings were shown to enhance retention of main-idea topics, $F(1, 111) = 7.51$, $p = .007$, but not for low-level detail (p. 220). Also, this effect only applied consistently to students with high prior knowledge, and not for the low prior knowledge ones. From these results, Wilhite concluded that:

…part of the beneficial effect of headings derives from their tendency to activate relevant schemas in the reader during the encoding of the passage material…(and not) by promoting the interrelating of concepts (or by influencing) retrieval by serving as cues at the time of test for subjects with low preexisting knowledge. (p. 223)

In other words, with recognition tasks (e.g., multiple-choice tests), learners with higher prior knowledge are able to make use of headings because the headings activate already-existing schemas. This, in turn, helps these students remember the main ideas of a passage, but does not necessarily help them remember more of the subordinate information. Learners with low prior knowledge do not benefit from the headings.

Headings effects on multiple-choice test performance by item difficulty. In a study that examined the retrieval effects of headings in a 43-item multiple-choice test to demarcate topic sections, Marcinkiewicz and Clariana (1997) found that headings significantly improved test performance for low ($es = 0.92$) and medium ($es = 0.44$) difficulty items but had no effect for high difficulty items ($es = 0.02$). This study involved 143 mostly-male, adult workers. No expository prose passages or lists were used as lesson material. The results of this study varied from a similar one (Townsend, Moore, Tuck, & Wilton, 1990) in which no significant differences for headings were found. Marcinkiewicz and Clariana (1997) attribute the difference primarily to the fact that they accounted for item difficulty level and the previous study had not. The results of this study are consistent with Wilhite (1988) in that easier test items (high prior knowledge) were influenced by signaling effects while difficult ones (low prior knowledge) were not.

Headings effects on free recall by familiarity. In two experiments examining the impact of headings on recall of prose, Lorch and Lorch (1996a) had 80 *Introduction to Psychology* undergraduates read a 2,400-word passage containing 12 topics and then complete a free recall task. Half of the topics were rated as familiar and half as unfamiliar. Although familiar topics were recalled more often than unfamiliar ones, the presence of headings in the passage improved overall recall of unfamiliar topics, $F(1, 60) = 6.02$, $p < .05$, but had no effect on familiar topics (p. 268). Consistent with other signaling research (Lorch & Lorch, 1995, 1996b), headings significantly improved recall of unfamiliar *main* topics, $F(1, 60) = 4.30$, $p < .05$, but had a slightly negative effect on recall of unfamiliar *subordinate* topics (p. 269).

Headings effects on text summarization by familiarity. In a second experiment involving 99 undergraduates and a 3,700-word reading passage containing 20 topics, Lorch and Lorch (1996a) found that headings improved recall on a summarization task for both familiar, $t(76) = 2.96$, $p < .05$, and unfamiliar topics, $t(76) = 3.68$, $p < .05$ (p. 272). Also, headings enhanced recall of topics that were only briefly discussed in the text, $t(76) = 4.99$, $p < .05$, but had little influence on topics that were discussed at length.

Do signaling effects vary by text length and complexity?

Signaling effects on recall of main and subordinate topics. Lorch et al. (1993) conducted two experiments testing the effects of signaling (i.e., headings, blank lines inserted between sections, overviews, and summaries) on recall of main and subordinate topics. In experiment one, subjects were 203 Introduction to Psychology students. Each was asked to read one of eight versions of a reading passage, which varied based on length (3,600 words with 24 idea units or 2,400 words with 12 idea units), complexity (2-level vs. 3-level topic structure), and signaling (all present vs. all absent). Signaling improved recall, $F(1, 194) = 67.00$, $p < .05$; signaling had greater impact on shorter text (2,400 words) than longer text (3,600 words), $F(1, 194) = 5.51$, $p < .05$; and signaling improved recall proportionately more for text with a simpler topic structure (2 levels) than text with a more complex organization (3 levels), $F(1, 194) = 112.73$, $p < .05$ (p. 283). To summarize, signals were effective encoding aids and were more effective when the text was short and/or simple (2 levels and 2,400 words) than when it was long and/or complex (3 levels and 3,600 words).

Signaling effects on recall of short and long discussion topics. In experiment two (Lorch et al., 1993), 82 subjects were asked to read a 2,000-word passage that varied based on signaling (all present or all absent) and the counter-balanced mixture of short and long discussions of topics. Signaling improved recall of subordinate topics that were embedded in shorter discussions of a topic, $F(1, 78) = 35.53$, $p < .05$ (p. 285). In fact, non-significant differences favoring the non-signaled condition were found in recall of topics in longer discussions. Similarly, they found that the non-signaled group was able to provide a greater amount of information about the topics they recalled than the signaled group, $F(1, 71) = 5.06$, $p <.05$ (p. 286). Lorch et al. surmised that while signals improve recall of topics, particularly with short passages, they may do so at the expense of stimulating recall of topic-related subordinate information. In other words, signals may help the reader remember the organization of the passage but not necessarily the substance of it.

Signaling effects on recall of main and subordinate topics. In two experiments involving 139 college students, Lorch and Lorch (1996b) asked students to read a 1,750-word passage that did or did not contain signaling devices (headings, overviews, and summaries) and complete a free recall task assessing recall of main and subordinate topics. The researchers hypothesized that signaling is most likely to have an effect on overall recall if the text is complex. In the first experiment, the simple topic structure (6 topics) was tested but non-significant differences were found. In the second experiment, which involved the more complex topic structure (12 topics), significant differences were found for recall of *main* topics, $t(66) = 2.71$, $p < .05$, and recall of *subordinate* topics, $t(66) = 2.42$, $p < .05$ (p. 42), favoring the signaled condition. Total recall was significantly improved as well, $t(66) = 2.03$, $p < .05$. Based on these results, signaling enhances overall recall when the reading passages are complex (12 topics) and do not necessarily enhance overall recall when the reading passages are simple enough for the student to encode the topic structure effectively without additional help. This is consistent with Spyrikakis and Standal's (1987) study, in which signaling enhanced recognition memory only when the text was long and complex.

On the surface, the findings from Lorch and Lorch's (1996b) study run counter to their earlier study (Lorch et al., 1993) in which signals were more effective with short, simple text. However, in the earlier study, Lorch et al. (1993) operationalized *complexity* by manipulating the number of topic *levels* (2 versus 3), while in this later study, they altered the number of topics themselves (6 versus 12). Thus, comparing the two findings is problematic. One could draw the conclusion from these two studies that signaling is most effective with text that has no more than two levels, approximately 12 topic and idea units, and approximately 2,400 words. Text with less than 12 topics may be too simple to benefit from signals and, conversely, text with much more than 12 topics may be too complex to benefit from signals.

Signaling effects on conditional recall. In a third experiment involving 120 college students, Lorch and Lorch (1996b) found that signals did not improve conditional recall (recall of subordinate topics when main topic is recalled), which rendered their impact on total recall non-significant. From this and their prior studies (Lorch & Lorch, 1995; Lorch et al., 1993), Lorch and Lorch (1996b) conclude that while signals "influence the distribution of ideas and the organization of recall," these gains "do not necessarily translate into better overall recall because signals do not consistently facilitate recall of subordinate content concerning a topic" (p. 45). Lorch and Lorch also found that half signals (i.e., including headings, overviews, and summaries for half of the topics but not for the other half) improved recall of main topics for the signaled topics but depressed recall of non-signaled topics. In other words, signals appear to improve recall of signaled topics at the expense of non-signaled ones. This is consistent with Lorch et al.'s (1993) earlier findings. It is also consistent with a study by Robinson and Hall (1941), conducted primarily to compare comprehension scores and reading rates for different types of text (art, geology, fiction, Canadian and Russian history). They found that headings had non-significant effects on comprehension test scores.

Do signaling effects improve with training?

Headings effects on recall, structural knowledge, and recognition. Brooks, Dansereau, Spurlin, and Holley (1983) examined the effects of three signaling conditions (embedded headings, outline, and embedded headings with outline) on three dependent measures: recall (essay test), structural knowledge (outlining test), and recognition (28-item multiple-choice test). In experiment one, involving 132 college students, Brooks et al. found that headings had a significant effect on recall, $F(1, 112) = 5.15$, $p < .02$, and on structural knowledge, $F(1, 112) = 6.41$, $p < .01$, but not on recognition (p. 296). Headings improve recall of main ideas and knowledge of the organizational structure of the text. These results also suggest that recognition memory retrieval tasks, such as those often involved in multiple-choice and matching tests, are less sensitive to signaling effects.

Headings effects on recall with and without training. In experiment two, involving 106 college students, Brooks et al. (1983) examined the effects of training students on the use of headings versus the effects of headings without training. Training significantly improved recall, $F(2, 102) = 4.53$, $p < .01$, and structural knowledge, $F(2, 102) = 19.90$, $p < .01$, but again not recognition (p. 299). Either the multiple-choice test was not able to detect any existing signaling effects or recognition tasks are immune to signaling effects, even with training.

Headings effects on free recall with and without training. Sanchez, Lorch, and Lorch (2000) also examined the impact of training on headings effects. In their study, 140 undergraduates read a 1,425-word text, then students in the training condition were instructed to form a mental outline of the text. This encouraged them to focus on the organization of the reading passage. Sanchez et al. (2000) found that students who neither had training nor headings had lower topic access scores (recall of main topics) and lower total recall scores than students in each of the other three conditions involving training and/or headings, $t(136) = 3.05$, $p < .05$ (p. 8). However, the effects of training and headings were essentially equal and did not interact to produce greater signaling effects. Access to subordinate information (i.e., conditional recall) was not affected by either training or the presence of headings. These results are consistent with prior studies (Lorch & Lorch, 1995, 1996a, 1996b) that found signaling effects for recall of organizational structure but did not find signaling effects for subordinate information.

Signaling effects on problem-solving

Signaling effects on free recall and problem-solving. In an experiment involving 58 tenth-grade college preparatory track students who were asked to read a 223-word passage of expository prose, Loman and Mayer (1983) found that headings and other signals caused students to modify their reading strategies, helping them recall more information and generate higher quality problem solutions. Students were given a reading passage that contained 43 idea units and, for those in the signaling condition, also contained preview sentences, underlined passage headings, and logical connective phrases. Idea units were blocked by type: conceptual, general, and primacy/recency (first idea/last idea). Loman and Mayer (1983) found significant differences between the signaled group (with headings) and the non-signaled group, $F(1, 56) = 9.80$, $p < .01$, and for the interaction of the three classes of idea units, $F(2, 112) = 315.01$, $p < .001$ (p. 406), suggesting that signals impact retention of different ideas disproportionately.

Signaling effects for remedial learners on free recall and problem-solving. In a second experiment involving 44 tenth graders who were designated as "remedial" and who read a 153-word passage containing 26 idea units, Loman and Mayer (1983) confirmed their earlier findings that the signaled groups remembered concepts and general idea units better than the non-signaled group. Conversely, they found that the non-signaled group retained more of the primacy/recency idea units. They concluded: "signaling was effective in asserting selective effects on *what is learned*" (p. 408). In other words, the paragraph headings, previews, and connective phrases drew the learner's attention to the conceptual information, lessening their dependency on primacy and recency.

In a study involving 32 college students and a 670-word passage dealing with the nitrogen cycle, Mayer, Dyck, and Cook (1984) examined the signaling effects of introductory definitions and headings on two criterion measures: a free recall test and a problem-solving test (10 essay or computational problems). The signals were written to communicate the relationship among the main idea units, rather than merely the topic structure. Responses from the free recall test were grouped into four categories: (1) recall of conceptual information, (2) recall of primacy information (first 10 idea units), (3) total recall, and (4) verbatim recall of information. Responses from the problem-solving problem were grouped into two categories: (1) near transfer problems, which required specific information to solve, and (2) far transfer problems, which required the learner to use information in a creative way.

Signaling effects of introductory definitions on recall of concepts versus verbatim information and far versus near problem-solving. In the first experiment, Mayer and his colleagues (1984) predicted that since the introductory definitions presented in the signaling condition were designed to communicate a relationship model, students would recall more conceptual information and perform better in problem-solving tasks at the expense of remembering verbatim and primacy information from the passage. The learners in the signaled condition did recall more conceptual information, $t(30) = 3.84$, $p < .01$, even thought total recall was not higher (p.1094). The signaled learners also recalled less primacy information, $t(30) = 2.20$, $p < .05$, and less verbatim information, $t(30) = 2.54$, $p < .05$ (p.1094). At the same time, they performed significantly better on the far transfer problem solving test than the non-signaled group, $t(30) = 3.35$, $p < .05$, confirming that introductory definitions facilitate more creative problem solving (p.1095. Signaling, however, had no effect on the near transfer problem solving items.

Signaling effects of headings and introductory paragraphs on recall of concepts versus verbatim information and problem-solving. In experiment two, Mayer et al. (1984) examined the signaling effects of headings and introductory paragraphs. The study involved 30 college students and the same reading materials used in experiment one. Consistent with the first experiment, recall of signaled concepts increased, while recall of primacy and verbatim information decreased. Signaling also improved problem solving, $t(28) = 2.10$, $p < .05$ (p. 1098). Signaling draws learners to main ideas and reduces their dependency on verbatim memorization. By conveying a meaningful organizational structure of the content, signaling enhances problem-solving.

Do verbal signals interact with non-verbal signals?

Mautone and Mayer (2001) ran three experiments to examine signaling effects on two criterion measures, a free recall test and a transfer test. They began by examining signaling effects on recall and transfer. Then, the text was converted to audio narration so that the effects of audio signaling could be examined. Lastly, the researchers compared the effects of audio signaling with visual signaling, i.e., graphic animation.

Signaling effects on recall and transfer. The first experiment involved 48 college students and a 487-word passage containing 7 main idea units dealing with the principles of airplane lift. The signaled version of the text passage was identical to the non-signaled version, except that the signaled version contained headings, a preview summary, and connecting words. The headings were bolded, underlined, and shown in a slightly larger (14-point) font, while the connecting words were boldfaced and italicized. The results indicated that while signaling did not significantly enhance recall, $t(46) = 0.478$, $p = .63$, it did significantly improve performance on a transfer test, $t(46) = 2.15$, $p < .05$ (p. 384).

Signaling effects with audio narration. In experiment two, which also involved 48 college students, Mautone and Mayer (2001) used the same 487-word passage from experiment one, but this time it was presented in the form of audio narration without text. In the signaled condition, the intonation of audio headings and connective words were varied and an audio preview summary was added. The results indicated that this time, signaling significantly enhanced both recall, $t(46) = 2.005$, $p < .05$, and transfer, $t(46) = 2.77$, $p < .01$ (p. 384).

Signaling effects with audio narration and animation. In experiment three, involving 86 college students, the same signaled and non-signaled audio narration were used, but this time, they were complemented with signaled and non-signaled animation. The animation was signaled using colored arrows as guides for selecting relevant parts of the illustration (as headings would); by using color to show the organization and relationships among the different parts of the illustration (as connective phrases would); and by including an icon that the student could click for a summary. The results suggest that signaled narration in combination with non-signaled animation enhanced recall, $F(1, 82) = 4.01$, $p < .05$ (p. 386), but that signaled narration was actually impaired by signaled animation. The transfer test results show a similar pattern favoring narration. When both signaled narration and signaled animation were used, students performed significantly better on the transfer test, $t(44) = 2.76$, $p < .01$. However, signaled narration significantly improved transfer, $F(1, 82)$, $p < .01$ (p. 386), while signaled animation appeared to have almost no impact.

Mautone and Mayer (2001) concluded that verbal signaling (text and audio narration) significantly enhanced transfer to new problem-solving situations, but had only slight to moderate effects on recall. Animation, a non-verbal signal, had little or no positive impact on recall *or* transfer. Their findings suggest that verbal signals are stronger than non-verbal ones, and that when used simultaneously, they can compete for attention and for short-term memory space.

Summary

When headings and other signals are absent from instructional text, learners must rely on other sources of information to identify the key points covered in the reading passage. These sources include the learner's prior knowledge as well as the serial positioning of topics (e.g., important topics first and less important topics later) and the amount of elaboration provided in the passage (e.g., important topics are explained more thoroughly). Signaling devices explicitly highlight the key points of a passage and communicate the structure of the text.

When are signals most effective? Signals do not necessarily increase the total amount of information recalled. Rather, they mainly alter the learner's reading/encoding strategy (Loman & Mayer, 1983). Several studies have shown that signaling significantly improves recall of main topic words and phrases (Brooks et al., 1983; Hartley et al., 1980; Krug et al., 1989; Loman & Mayer, 1983), with both cued and free recall tasks. Signals, however, do not necessarily improve recall of subordinate information (Lorch & Lorch, 1996a). In fact, headings and other signals shift the focus of attention during encoding to the signaled information and away from the non-signaled information (Krug et al., 1989; Lorch & Lorch, 1996b; Robinson & Hall, 1941). Although signaling may not improve recall of subordinate information, it does enhance transfer of knowledge to new situations, thus aiding in problem-solving tasks (Mautone & Mayer, 2001; Mayer et al., 1984). Signaling is an appropriate strategy when recall of the cued topics is important and when the key topics are communicated to facilitate problem-solving.

Multiple-choice, matching, and other recognition item formats appear to be less sensitive to signaling effects compared to cued recall and free recall tasks (Christensen & Stordahl, 1955; Marcinkiewicz & Clariana, 1997; Spyridakis & Standal, 1987). Smith's (1986) outshining hypothesis posits that environmental cues enhance memory retrieval when the to-be-remembered information is difficult and do *not* when it is easy to remember. Recognition tasks are generally easier than recall tasks, and so this may explain why signaling is less effective with multiple-choice tests than it is with recall tests. The difference may also be due to ceiling effects (Spyridakis & Standal, 1987), which are more likely to influence easier test formats.

Statement headings and question headings are equally effective (Hartley et al., 1980). Headings are not any more effective than other signaling devices (Lorch & Lorch, 1995). However, headings are highly effective when grouped with other verbal signals (Brooks et al., 1983; Hartley et al., 1980; Loman &

Mayer, 1983; Lorch & Lorch, 1996a, 1996b; Lorch et al., 1993) but are not more effective when combined with non-verbal signals (Mautone & Mayer, 2001).

Signaling effects are strengthened when the information is cued at retrieval (Lorch & Lorch, 1995). Signals are e specially helpful when the text is approximately 2,400 words in length and has approximately 12 topics with no more than two levels (Lorch & Lorch, 1996b; Lorch et al., 1993). Text that is too short and simple or too long and difficult may not benefit from signaling effects.

Implications and Conclusion

Instructional designers and educators commonly use headings and other signals when writing instructional text for paper-based and screen-based materials, regardless of the instructional objectives. The research suggests that headings and other signals can be powerful textual cues that alter the encoding strategy of learners. While headings can promote recall of main topics and enhance problem-solving, they are not helpful in promoting reading comprehension. In fact, signals tend to cause learners to remember key ideas at the expense of subordinate information that may be crucial to understanding. This implies that writers of instructional text should consider the kinds of instructional objectives sought before deciding whether or not signaling will be helpful or counter-productive.

If headings are being used for the purpose of enhancing recall, they will be most effective when used in conjunction with other verbal signaling devices, such as introductions, summaries, transition statements, and titles. Headings will be most useful when the text is moderately long and complex, and when comprehension of the subordinate information is less important than understanding of the main points.

Much of the signals research involves paper-based instructional materials but few involve screen-based materials. In online venues, instructional text can assume complex forms, with text hyperlinked to subordinate or related content, so the text is not always sequentially ordered, and elaboration is dependent on the user's desire to explore by clicking to reveal deeper layers of information. In such settings, would verbal signaling devices be more helpful or less helpful in promoting recall, comprehension, and problem-solving? Given the increasing use of non-verbal cues, such as icons, graphics, animations, and color, do they compete with or else complement verbal signals in different virtual settings? Given the proliferation of computer-based and web-based courseware, online distance learning, and electronic textbooks, future signals research should reevaluate the effects of headings and other signals in the digital medium to provide designers clearer answers to such questions.

Headings are valuable signaling devices that have been shown to promote recall of main ideas, knowledge of organizational structure, and problem-solving. They draw the learner's attention to the key topics in a reading passage. These key topics form the topography of the text based on the organizational structure of the passage. As such, they implicitly contain relationships and other structural information. Tasks and skills that benefit from this improved structure of the content are therefore likely to benefit from signals. However, it is critical to recognize that these gains are made at the expense of memory for subordinate information.

References

Brooks, L. W., Dansereau, D. F., Spurlin, J. E., & Holley, C. D. (1983). Effects of headings on text processing. *Journal of Educational Psychology, 75*, 292-302.

Christensen, C. M., & Stordahl, K. E. (1955). The effect of organizational aids on comprehension and retention. *Journal of Educational Psychology, 46*(2), 65-74.

Hartley, J., Kenely, J., Owen, G., & Trueman, M. (1980). The effect of headings on children's recall from prose text. *British Journal of Educational Psychology, 50*, 304-307.

Hartley, J., & Trueman, M. (1985). A research strategy for text designers: The role of headings. *Instructional Science, 14*, 99-155.

Krug, D., George, B., Hannon, A., & Glover, J. A. (1989). The effect of outlines and headings on readers' recall of text. *Contemporary Educational Psychology, 14*, 111-123.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology, 75*(3), 402-412.

Lorch, R. F., Jr., & Lorch, E. P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology, 87*(4), 537-544.

Lorch, R. F., Jr., & Lorch, E. P. (1996a). Effects of headings on text recall and summarization.

Contemporary Educational Psychology, 21, 261-278.

Lorch, R. F., Jr., & Lorch, E. P. (1996b). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology, 88*(1), 38-48.

Lorch, R. F., Jr., Lorch, E. P., & Inman, W. E. (1993). Effects of signaling topic structure on text recall. *Journal of Educational Psychology, 85*(2), 281-290.

Lorch, R. F., Jr., Lorch, E. P., Ritchey, K., McGovern, L., & Coleman, D. (2001). Effects of headings on text summarization. *Contemporary Educational Psychology, 26*, 171-191.

Marcinkiewicz, H. R., & Clariana, R. B. (1997). The performance effects of headings within multiple-choice tests. *The British Psychological Society, 67*, 111-117.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*(2), 377-389.

Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraining and signaling. *Journal of Educational Psychology, 76*(6), 1089-1105.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory*. New York, NY: Elsevier Science.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 1*, 72-101.

Meyer, B. J. F., & Rice, E. (1982). The interaction of reader strategies and the organization of text. *Text, 2*, 155-192.

Robinson, F. P., & Hall, P. (1941). Studies of high-level reading abilities. *Journal of Educational Psychology, 32*, 241-252.

Sanchez, R. P., Lorch, E. P., & Lorch, R. F., Jr. (2000). Effects of headings on text processing strategies. *Contemporary Educational Psychology*.

Smith, S. M. (1986). Environmental context-dependent recognition memory using a short-term memory task for input. *Memory & Cognition, 14*, 347-354.

Spyridakis, J. H., & Standal, T. C. (1987). Signals in expository prose: Effects on reading. *Reading Research Quarterly, 22*(3), 285-295.

Townsend, M. A. R., Moore, D. W., Tuck, B. F., & Wilton, K. M. (1990). Headings within multiple-choice tests as facilitators of test performance. *British Journal of Educational Psychology, 60*, 153-159.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.

Wilhite, S. C. (1988). Reading for a multiple-choice test: Headings as schema activators. *Journal of Reading Behavior, 20*(3), 215-228.

Educators' Perceptual Barriers: Examining the Diffusion of Educational Technology and Education Reform

LeAnne K. Robinson
Western Washington University

Abstract

The study examines educator's perceived barriers to technology integration and the relationship to education reform. Educators and administrators from four elementary schools located within the same educational service district in western Washington were interviewed in their classrooms and offices during a three month period. The schools differed in size, location and social economic status and reported variances in their Washington Assessment of Student Learning (WASL) scores. The study found that while all of the schools reported similar barriers to the use of educational technology, distinct differences appeared between those schools that had done long range planning during the reform process and those that had not. Specifically, the two schools that coordinated curricula, performance standards and a variety of assessment tools while simultaneously allowing teachers the flexibility to alter the curricula as needed, were more likely to state personal responsibility for student learning and were more likely to have overcome barriers to the use of technology. Three assertions were generated: (a) When education reform efforts include the integration of curriculum, assessment and classroom instruction, teachers are more likely to take personal responsibility for student learning; (b) Teachers will take more responsibility for technology use when technology integration is not separate from curricula and/or reform efforts; (c) In order for technology to be used in a school, access to technology needs to be made for all. A model for conceptualizing technology integration is presented. Discussion includes the need to create schools where teachers feel control over both the classroom and the technology, are able to take responsibility for student learning and the use of technology as well as have a sense of accomplishment.

Background

With the No Child Left Behind initiative and the corresponding 'Enhancing Education Through Technology Program' that provides funding to increase student achievement through the use of technology in schools, it is apparent that both education reform and the integration of technology into classrooms continue to be valued and supported. Nearly every state has adopted or has developed assessments that align with standard-based reform efforts (Stechner & Chun, 2001). Often technology reform and education reform have paralleled each other as opposed to being incorporated (Peters, 2000). This means that in many instances the purchasing of computing tools and related technology, as well as plans for staff development, were not coordinated with school reform, nor with the concurrent examination of curricula and instructional processes. When both education reform and technology integration have been fully combined into curricular reform, including the examination of pedagogy, positive results have been found for students (Bain & Smith, 2000).

Even with all of the pressure to integrate technology into the curriculum, the presence and accessibility of computers in the schools has not shown that the technology is being used by educators or that students actually can or do use it (Cuban, Kirkpatrick & Peck, 2002; Kalkowski, 2001). Although they are accessible, computers have not transformed the practices of a majority of teachers (Becker, 2000; Labbo & Reinking, 1999) and Willis, Thompson, and Sadera (1999) have pointed out that integration of computers into the classroom has actually been a slow process.

The literature identifies many reasons for the lack of integration. These include inappropriate training or inservice (Kay, 1996; Maor, 1999); collegial jealousy or pre-defined roles (Sherry & Billig, 2002; Reinking & Watkins, 2000; Wood, 2000); lack of appropriate or relevant software instruction (Becker, 2000; Rockman, 2000; Rogers, 2000; Ruberg, 1993; Sia, 1992); teachers failing to find the relevance of technology use or applications to classroom practices (Maor, 1999; Sherry, Billig, Tavalin, & Gibson, 2000; Rogers, 2000); as well as teachers having a lack of space and time within the curriculum (Cuban, Kirkpatrick, & Peck, 2001; James, Lamb, Bailey, & Householder, 2000).

Individual rationales for a lack of technology integration have been referenced as "primary and secondary barriers" (Judson & Sawada, 2000; Prater, 2001; Ertmer, Addison, Lane, Ross, & Woods, 1999).

Both primary and secondary barriers are explained as being both intrinsic and extrinsic (Ertmer, Addison, Lane, Ross, & Woods, 1999). Primary barriers include: lack of access to computers and software, insufficient time to plan instruction, and inadequate technical and administrative support. Secondary barriers include: beliefs about teaching, beliefs about computers, established classroom practices, and unwillingness to change (Ertmer, Addison, Lane, Ross, & Woods, 1999).

The diffusion of innovations is the study of the process by which the use of a perceived new idea, practice or object is adopted within a given social system (Rogers, 1995). The study of the diffusion of innovations is present in many research traditions, including: anthropology, marketing, geography, communication and education. Within the overall research arena, less than 10% of the studies of innovations have been conducted in education (Rogers, 1995). While Rogers (1995) provides a generic model of the process of the adoption of an innovation, recent research is showing that alternative models may be more applicable to school systems. These models specifically identify educational technology as the innovation being studied, thus the phrase 'diffusion of educational technology' can be used in place of 'diffusion of innovations'.

The diffusion of educational technology models that have been recently presented in the literature are non-linear, implying that many factors are involved in the adoption (or lack of adoption) of educational technology. These include both Rogers (2000) and Sherry, Billig, Tavalin, & Gibson (2000) as well as one presented by James, Lamb, Bailey, and Householder (2000). A holistic model has been offered where relevance, principal leadership, and intrinsic and extrinsic factors, including pedagogical beliefs, practices and exposure to innovations, are considered (Dooley,1999). These models attempt to identify more specifically factors that will lead to the adoption of an innovation in a school setting. However, the cited models tend to be based on single technologies, as opposed to the full integration of technology. The models do not address the multiple barriers that are cited within the literature. In order to overcome a barrier to integration, it was the assertion of the researcher that it is necessary to understand how technology diffused and how the diffusion relates to various barriers.

In the spring of 2001, the researcher conducted a pilot study at a K-6 elementary school to determine how teachers in this particular setting made decisions regarding media selection. While the majority of staff verbally supported the use of computing tools, few of them used educational technology and instead tended to blame others or cite circumstances that they felt were out of their control for a lack of use. The common theme that emerged from all of the interviews was that regardless of the reason for failing to integrate technology, the teachers in the building were making the majority of decisions based on a highly structured reading program that had been adopted during a school reform process. All related the lack of technology use to some aspect of the school's focus on educational reform and to their concerns related to teacher accountability and the Washington Assessment of Student Learning (WASL). The results of the pilot study provided a foundation for a study examining the relationship between education reform, educational technology and educators' perceived barriers to technology integration.

Research Questions

The following research questions guided this study:

1. What are teachers' perceptions of barriers to the use of educational technology/technology integration and how do the barriers connect to education reform?
2. What are the connections between perceived barriers to the responsibility for computing technology?

A four-stage model proposed by Clark and Estes (1999) for conducting research in the development of authentic technologies provided the theoretical underpinnings for this study. Authentic technologies are "educational solutions resulting from systematic analysis that identifies the problem being solved, selects and translates appropriate, well designed research and applies it to design culturally appropriate educational solutions", (p. 243). For this study, authentic technology included teaching strategies and processes, not merely computing tools. This research is based in the first stage, or Descriptive Scientific Research Stage, where the defining of construct definition and hypotheses generation are the key goals.

Site Selection

Four sites from separate districts were purposively selected from the same Educational Service

District in western state and a stratified sample was used (Patton, 1990), meaning that the four sites represented four different subgroups for comparison. Two of the schools were located in rural settings. The schools were selected based on self reported variances in their WASL scores, with one of the schools having exceeded the state average in percentage of students passing. Demographics are reported in Table 1. Initial access was granted through a key individual at each site and snowball sampling (Patton, 1990) was used.

Twenty separate individuals were interviewed between March and May of 2002. Interviewees included three administrators, one administrative intern, three reading specialists, thirteen classroom teachers, a counselor and a P.E. teacher. Each interview lasted between forty-five minutes and an hour and a half. All of the interviews occurred within the individual teacher's classroom or the administrator's office at the school site. Following the transcriptions of the interviews and during the analysis, the initial contact person at each school was available to answer specific questions via phone and email.

Table 1 *Enrollment and WASL Scores of Selected Sites, 2001*

School	Location	Enrollment	% of Free and Reduced Lunch	% Passing Reading WASL (66.1 state)	% Passing Math WASL (43.9 state)	% Passing Writing WASL (43.5 state)	% Passing Listening WASL (72.4 state)
Wrangle	Urban	370 K-6	56.8	48.8	23.3	27.9	58.1
Sandal Creek	Suburb	444 K-4	10.3	52.7	28	43	71
East Lake	Rural	263 K-6	34.4	55.8	30.2	23.3	72.1
Wood-land	Rural	304 K-6	44.4	73.2	39	43.9	80.5

The International Society for Technology in Education (ISTE) has developed frameworks that provide progressive descriptions of teachers who do not integrate technology to those teachers that fully integrate technology (ISTE Homepage, 1999). These frameworks were converted into a survey and each teacher was asked to identify his/her current level of technology integration. Notes were taken during the interviews and during observations. Comparisons were then made between the teachers reported level of integration, information gathered during the interview, and the researcher's observations in order to make a determination of the accuracy between reported and actual levels of technology integration. Field notes and a methodological log were kept. Each interview was recorded and transcribed verbatim.

Data Analysis

Each interview was read multiple times and the researcher looked for themes within each broad category. A list of themes was generated and the interviews were again re-read and selective sections of the interviews that corresponded to the potential themes were labeled. For the teachers, the following themes were identified: primary barriers, secondary barriers, school climate, favorite parts of teaching, frustration/needs, teaching practices, accountability and technology specific responses. These themes then served as categories, and the responses were read again and sub-categories were looked for within each theme. The researcher then created an organizational chart and returned to the interviews, coding the specific parts of each interview that related to each specific category.

The interviews were photocopied and each school was given a unique color so that the schools could be identified within the organizational chart. Additional color coding was used to identity the level of technology integration that was reported on the technology integration survey. The interview selections were then cut and taped onto the charts. As the interviews were being sorted, the researcher further separated responses by grouping the responses of those who had high integration scores on the technology integration survey and those that scored lower. In addition, the researcher separated the responses to both primary and secondary barriers into two groups: those who saw an identified barrier as an obstacle that they could not or would not overcome and those that were attempting to or had overcome the identified barrier.

The researcher also looked for patterns within and between schools, between those who rated themselves high on the integration survey and those who rated themselves as lower on the survey. Initial conclusions were drawn, but, to be sure, an individual profile for each school was developed. The profiles included a summary of the schools responses to each theme as well three years of WASL and the Iowa Test of Basic Skills (ITBS) trends, demographics, and teachers interviewed. These brief profiles were then compared to each other to assist in confirming the researcher's initial findings from the organizational charts.

Interpretation

Assessment, Curriculum and Teacher Responsibility

Clear distinctions appeared between the four schools. All of the barriers to integrating technology were cited at all of the schools. What differed was how teachers responded to these barriers and how the schools viewed assessment and curricula (See Table 2).

Table 2 *Comparison of Assessments Curriculum*

	Schools			
	Wrangle	East Lake	Sandal Creek	Woodland
Primary Assessments				
Classroom		X	X	X
School	X		X	X
District			X	X
State	X	X	X	X
Role of Curricula				
Highly Structured	X		X	X
Adopted Curricula and District Frameworks with Flexibility to Adjust			X	X
No adopted Curricula and District Frameworks		X		
Teachers State Responsibility for Student Learning			X	X
Teachers State Responsibility for Technology Integration	X*		X	X
Teachers State Plans for Overcoming Barriers	X*		X	X

* Indicates a single teacher

From the analysis, three assertions were generated: (a) When education reform efforts include the integration of curriculum, assessment and classroom instruction, teachers are more likely to take personal responsibility for student learning; (b) Teachers will take more responsibility for technology use when technology integration is not separate from curricula and/or reform efforts; (c) In order for technology to be used in a school, access to technology needs to be made for all.

The first two distinctions dealt directly with personal responsibility. At two of the schools, Woodland and East Lake, teachers made statements related to being responsible for both student learning and for the use, or lack of use, of technology. These were the two schools that had made efforts to connect curriculum, assessment, and best practices with an overall plan for school improvement during the education reform processes. The districts had adopted curriculum, attempted to align the curriculum with state benchmarks and standards and used multiple types of assessment, including formal and informal classroom, district and state assessments. Teachers were expected to utilize the curriculum but were simultaneously given flexibility to make adjustments or alterations based on the teacher's professional knowledge and the needs of the students.

> …and if all of that work has been done [the curriculum process] the curriculum is critical; how its adjusted and adapted and delivered to the students is totally a professional obligation and responsibility of that classroom teacher, but they need to be able to know the curriculum well enough to how to adapt it. I would encourage all teachers the first year to use the curriculum closely….they are going to have to do some adapting and they'll probably have to work with the special education teachers …but over time, as they become familiar with the materials, they know how to make those adaptations and differentiate within the classroom setting. Again, we don't teach to the book, we teach kids and you have to figure out where they are and it not just where they are for this year, it is where they are today……we check every day on how they are doing. *(Principal, Woodland, May 10, 2002)*

The remaining two schools, Wrangle and Sandal Creek, had taken opposite approaches to education reform. Wrangle had students with low reading scores, and had responded by adopting a highly structured highly scripted reading program that impacted all of the decisions that were made:

> It is a structured program that we are working with….They have scripted materials for me to use. *(Second Grade Teacher, Wrangle, March 28, 2002)*

Whereas at Sandal Creek, very little structure existed:

> Actually I am sitting on a committee right now that is trying to purchase a curriculum. We are all frustrated with the hodgepodge.
> *(Third Grade Teacher, Sandal Creek, May 8, 2002)*

At each of the schools, teachers cited the same barriers to technology integration. However, it was apparent that theachers at the two schools who had coordinated school reform efforts responded differently to the barriers. These teachers would talk about the barriers that existed and then describe how he/she had overcome or attempted to overcome the barrier that had been presented. For example, the following was a typical comment from teachers at East Lake who faced the lack of newer computing tools:

> I saw them in the hallway. Two of them are old Apples, but I have a lot of Mac materials, so that a lot of drill and practice things and I mean those others just came in this year, they were from the middle school and they got new ones so they were in our hallway. I think there were twenty so we each got a couple more, so that is how we got three or four because nobody wanted them….I think it is hard because we don't have technology for the whole school. *(Second/Third Grade Teacher, East Lake, May 6, 2002)*

As opposed to the same barrier at Sandal Creek, where the older computers were not wanted:

> ….they [the District technology committee] have a plan that would make it so that you have a new computer in every school every three years, but that doesn't mean we'd get a new computer, that means probably the high school would get the new computer, the junior high would get their old computers, so that there is a new machine. It is new to the school but not necessarily a new machine…..…we've talked about a computer lab…[in regards to the district technology plan]…so, I don't want any more hand me downs, just take them away. I don't need them. *(Third/Fourth Grade Teacher, Sandal Creek, May 8, 2002)*

Teachers at Woodland and East Lake, made statements related to taking responsibility for student learning, as one teacher from Woodland put it, "It's the teacher's responsibility". They also made statements about being personally responsible for the use, or lack of use of technology. For example, one teacher at Woodland commented on her observation that it was actually her problem for failing to use technology, recognizing that she had used 'lack of access' as a scapegoat:

It is definitely not what we have now, you know, last year I could have said we don't have great computers. But I think it is just a comfort level for me and I think I need to get in and experiment to how to use it myself and then be able to expect my kids do that, so I think it is something that will happen, maybe next year….*(Third Grade Teacher, Woodland, May 10, 2002)*

This type of response, or statement of personal responsibility, was not apparent at either Sandal Creek or Wrangle. When barriers to technology were cited, only one teacher stated responsibility for the use of technology. None of the teachers made statements about being responsible for student learning.

> …because people aren't telling us what the research base is, so I'm just finding whatever's there and that's what I will use, but I'm not using it to teach kids with, I'm using it as like a reward for having completed whatever things they've done in the classroom, so I'm not thinking of it as, in terms of curriculum. It is more like…it is not part of our curriculum, it is something extra for kids who are finished and need an extension of whatever. *(Second Grade Teacher, Wrangle, May 7, 2002)*

> I know there needs to be some sort of accountability, but I wish that this accountability, the accountability issues and the ed. reform issues were less centered on the teachers and the need for schools to make some arbitrary improvement and the politicians who actually come up with this stuff would actually point their finger at where the real issue is, which is in the home. *(Sixth Grade Teacher, Wrangle, March 8, 2002)*

Examining Single Technology Rich Classrooms

The concept of homophiliy (Rogers, 1995) suggests that the adoption of a new technology is more likely to occur if someone similar, such as an equal colleague, introduces it or is successful at using the new technology. It was interesting to note that Wrangle school had more computers per student than any of the other schools. Three of the teachers had received a total of four large grants and every teacher had at least one computer in his/her classroom. Of all of those interviewed, Wrangle also had a teacher that scored higher than any other did on the integration survey. However, Wrangle had less technology use as measured by the Technology Integration Scale. The large grants received by one specific teacher seemed to have isolated her. The technology provided by the grant stayed within the individual classrooms, although the grant recipients did collaborate with each other. On the other hand, Woodland had received a large grant where the technology moved with the students. Following the initial grant year, the student took the technology, in this case portable word processors, with them to the next grade. The new teachers then began utilizing the technology because the students were accustomed to it and it was still available to them.

One teacher reflected on how this model impacted other teacher's willingness to integrate technology:

> …well, it began, this is its second year…who is now our district librarian…I stepped in the second half. I said, 'Sure I'll try that', thinking, 'Okay, my technology skills are lack"…and the training was there and it didn't cost me a dime….It was four 5th graders but the kids who were 5th graders last year were able to, the LD kids were able to hook up with their laptops to 6th grade and use them in 6th grade, so it…we have two new 6th grade teachers and they, its kind of branched out and reached into their classrooms. *(Fifth Grade Teacher, Woodland, May 10, 2002)*

Wood (2000), found that teachers who were next to a grant recipient often were jealous and those with the technology were often discouraged from sharing, isolating them from their peers. This was seen at Wrangle. It was found that other teachers were often angered by the technology grant recipients who, every year, collaborated in the creation of a play. Two classrooms were involved in the development of the play and the integration of technology throughout was apparent: brochures, advertising, film editing, lights, music were all developed utilizing technology. One of the teachers who was interviewed and had observed the play, asked the researcher to figure out, "How, exactly, does that fit with curriculum"?

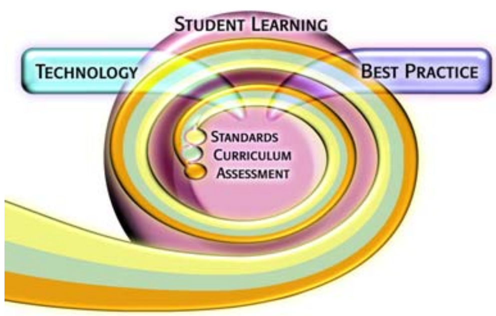
Discussion

Two diffusion models of educational technology identified that the point of rejection of a new technology occurred when teachers failed to see relevance for the learner (Sherry, Bilig, Tavalin & Gibson, 2000; Rogers, 2000). Upon completion of this study it seems necessary to further define the diffusion of educational technology models so that they encompass more than one technology and provide direction in the overall process of adopting multiple technologies that enhance all aspects of the curricula, not just the adoption of a single technology.

In reality, schools cannot be centered solely on technology, or just assessment, or curricula. Schools that become too focused on one area can neglect the others. In order to better understand how technology can be supported, the school needs to be viewed from the perspective of student learning first. What are the factors that directly influence and impact student learning? Such things as the relationship of curricula, assessment, accountability, leadership style, resources and the individual needs of the teacher need to be examined in a global context. Technology integration fits in to many of the categories, but technology alone won't guarantee student learning. It is simply one factor that may contribute to assisting a student in meeting learning goals.

The integration of technology into today's classroom needs to be viewed as an integral piece of education reform. Although the literature often uses the phrase "integrate technology" to imply that technology is to be part of multiple areas of education, including curriculum, assessment and instruction, it is often presented as the focal point or treated as a separate component. ISTE has presented the necessary conditions that need to be present in order for technology to be integrated (ISTE, 2002). Some of the influential factors include a shared vision, access, and skilled educators. Although these are important, the focal point starts with technology, not necessarily student learning, although there is the assumption that the use of technology will lead to greater student achievement. Perhaps a more appropriate phrase would be to, "Enhance teaching and learning through the effective use of technology". Student Learning should be the focal point for decisions regarding standards, curriculum and assessment. Standards, curriculum and the assessment should be tightly connected and directly related to student learning. The use of technology should be viewed as an integral and connected component for decisions regarding standards, curriculum and assessment. Best practice, like technology, is not separate but an integral component of all practices (Figure 1).

Figure 1 Conceptualization of the Role of Technology



Focusing on Student Learning and Supporting Teachers

Teachers at schools throughout the United States are frustrated and burn-out is high. Even if administrators are able to re-conceptualize the role of technology in a holistic fashion, teachers need to have barriers addressed on an individual basis and the need for support remains. One method for identifying how teachers can be supported may lie in the examination of motivational principles and theories from other areas that relate to human and computer interaction. For example, there are three fundamental principles in interface design that need to be present for a computing program to be used (Schneiderman, 1998). These principles seem appropriate to consider when asking if the conditions within a school are present to promote student learning through the effective use of technology. These principles are: (a) user-centered; (b) promote responsibility; and (c) provide a sense of accomplishment.

User-Centered or User-Control

The premise is that users have a strong desire to be in control or gain mastery of the system. In the context of a classroom, administrator's need to ask, 'Do the teachers in this building feel in control of his/her classroom'? Being told what to teach, when to teach, what to assess and how to assess by someone outside of the classroom lessons personal control. In the schools studied, the idea of control came up repeatedly. At Wrangle, teachers felt they had little control in terms of curricula, assessment and media selection. At Sandal Creek, teachers were given so much freedom that they felt out of control. All of the teachers commented on frustration with the lack of guidelines. In contrast were Woodland and East Lake where teachers were provided with curricula and district frameworks. Administrators verbally supported teachers in making changes to the curricula to meet the needs of the students. Although teachers at these two schools were not without frustrations, they were able to focus on students, as opposed to focusing on problems related to 'the administration'. Woodland and East Lake teachers also felt empowered to make decisions about the use, or lack of use, of technology. As with the theory of interface design, without control by the potential user, responsibility, the second premise, will not be taken.

Promote Responsibility

Once a user has control, the user is able to accept responsibility for his or her actions. The question for an administrator is then, 'Do the teachers take responsibility for what is occurring in the classroom'? and if the answer is 'no', the second question would be to ask not necessarily 'Why', but 'How can I support the teachers in a way that makes them feel in control'? At Sandal Creek, teachers blamed the school district's technology committee for a poor plan and blamed the school district for not providing curricular materials. There were virtually no stated connections between assessment and instruction. At Wrangle, teachers wanted someone else to tell them 'what the research was' and the principal even deferred the selection of school software to the librarian, stating that it was 'easier to sign the purchase order'. The teachers at Wrangle were being held accountable for test scores, but the agents of control were outside of the classroom; a reading coordinator and individuals from the reading program itself provided the assessments and directed the changes to instruction or to the configuration of the reading groups, not necessarily the teachers.

Provide a Sense of Accomplishment

The sense of accomplishment only comes from having the feeling of control and then responsibility. Within interface design, if a user does not have a sense of accomplishment, they will not continue to use a program. In a classroom, teachers need to feel accomplishment and need to feel that their efforts lead to student learning. In both of the educational technology diffusion models, teachers ceased to use an innovation if he/she did not see the connection to student learning (Sherry, Bilig, Tavalin & Gibson, 2000; Rogers, 2000). Having a lack of control, or a perceived inability to overcome barriers to technology integration, leads to a lack of responsibility and no accomplishment can be felt. Teachers will not attempt to overcome barriers. Accomplishment for teachers at Wrangle was directly related to performance on the WASL. This was not the case at Woodland or East Lake, where the WASL as viewed as a minor snapshot of student and teacher achievement. Teachers at these schools and the administrators placed more importance on the gains that individual students made on the classroom and school assessments.

Implications for Professional Practice

In order for teachers to overcome barriers to technology integration, teachers need to feel in control both of the classroom and the available technology, be able to take responsibility and have a sense of accomplishment. Although these three premises can be applied to the use of a single technological innovation, they should be applied to technology integration as a whole both within the classroom and within the context of an entire school system. Too many schools are taking away the ability for teachers to be responsible. Teachers are being given scripted curricula, are told what to teach and which assessments must be given, as opposed to being given the freedom to make appropriate choices about instruction and appropriate assessment.

Several recommendations can be made for school improvement and restructuring.
First, combine technology with other reform efforts. Make the connection between the use of certain

technologies or best practice with continued student achievement. This can take many forms; projects for increasing student learning, specific software applications for student use, appropriate assessment, and continued classroom management. Isolating technology, or any other component of reform, such as curricula, creates frustrations and limits vision.

Secondly, schools need to have the ability to overcome multiple primary barriers. Focusing on one barrier, such as access alone, neglects the others. Teachers are unique, have differing skills and need different opportunities for growth. Meeting one need does not guarantee integration: other barriers will arise and the ability to address them needs to be available for all. This includes not limiting materials and resources provided by a grant to a single classroom. This can isolate a teacher from his/her peers and create potential school climate issues. If a teacher does receive a grant, plans for future growth for other teachers within the building need to be thought through. This is the same for the students. What about those students who spend a year in a technology rich classroom followed by several years in a technology poor classroom, or those students at the same grade, because of placement in a certain classroom are then denied access to computing tools? Solutions to the one technology classroom may include team teaching, or cross-age projects.

Thirdly, long range planning needs to be done for all school restructuring. Although in Washington State schools are required to submit building goals for the following year, these do not always appear to be tied to long-range plans. School goals need to be created in a holistic context that includes the close scrutiny of curriculum, assessment practices, how teachers are matching classroom practices with established frameworks and how teachers are implementing best practices in instruction. Building goals should be developed followed with the examination of individual teacher needs. It simply isn't enough to set goals without taking into consideration different barriers that individual teachers may face.

Conclusions

The purpose of this study was to examine the relationship between education reform and teachers' perceptual barriers to the integration of technology. The four schools, all unique, had the same perceptual barriers to the integration of technology. At two of the schools where standards, curricula, and assessment were coordinated, teachers were more likely to state personal responsibility for student learning and the use of technology. The field of instructional technology is only beginning to develop models of diffusion that explain how technology is adopted by teachers. From this study, it appears necessary to begin to look at barriers not individually, but in the context of an entire school system, while further exploring how theories of motivation can positively influence both professional development and school reform efforts that incorporate the effective use of technology.

References

Bain, A., & Smith, D. (2000). Technology enabling school reform. *T.H.E. Journal, 28*(3), 90-97.

Becker, H. J. (2000). Findings from the teaching, learning, and computing survey: Is Larry Cuban right? *Education Policy Analysis Archives, 8*(51).

Clark, R.E., & Estes, F. (1999). New directions: How to develop "authentic technologies". In R.E. Clark
(Ed.), Learning from media: Arguments, analysis, and evidence (pp.241-262). Greenwich, CT: Information Age Publishing.

Cuban, L., Kirkpatrick, H., and Peck, C. (2001). High access and low use of technologies in high school classrooms: Explaining an apparent paradox. *American Educational Research Journal, 38*(4), 813-834.

Dooley, K.E. (1999). Towards a holistic model for the diffusion of educational technologies: An integrative review of educational innovation studies. *Educational Technology and Society 2*(4). Retrieved from the World Wide Web October 17, 2001.http://ifets.ieee.org/periodical/vol_4_99/kim_dooley.html

Ertmer, R.A., Addison, P., Lane, M., Ross, E. & Woods, D. (1999). Examining teachers' beliefs about the role of technology in the elementary classroom. *Journal of Research on Computing in Education, 32*(1), 54-62.

ISTE (2000). National Educational Technology Standards for Teachers. United States and Canada: International Society for Technology in Education.

ISTE Homepage. (1999). Online supplement: Levels of technology implementation framework. Learning and Leading with Technology 26(8). Retrieved from the World Wide Web April 16, 2002.

http://www.iste.org/L&L/archive/vol26/no8/supplements/moersch/table2.html
James, R.K., Lamb, C.E., Bailey, M.A., & Householder, D.L. (2000). Integrating science, mathematics,
and technology in middle school technology-rich environments: A study of implementation and change. *School Science and Mathematics 100*(1), 27-35.

Judson, E., & Sawada, D. (2000). Examining the effects of a reformed junior high school science class on students' math achievement. *School Science and Mathematics 100*(8), 419-425.

Kalkowski, M.A. (2001) Focus on learning and technology. *Communication: Journalism Education Today, 34*(4), 19-22, 25, 27.

Kay, A. (1996). Revealing the elephant: The use and misuse of computers in education. *Educom Review 31*(4). Retrieved from the World Wide Web November 5, 2001. http://www.educause.edu/pub/er/reviewArticles/31422.htm.

Labbo, L.D., & Reinking, D. (1999). Negotiating the multiple realities of technology in literacy research and instruction. *Reading Research Quarterly 34*(4), 478-492.

Maor, D. (1999). Teachers-as-learners: The role of multimedia professional development program in changing classroom practice. *Australian Science Teachers Journal, 45*(3), 45-50.

Patton, M.Q. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park: Sage Publications.

Peters, L. (2000). Joining forces: A third millennial challenge: Harness the power of educational technology to advance the standards movement. *T.H.E. Journal, 28*(2), 94-102.

Reinking, D. & Watkins, J. (2000). A formative experiment investigating the use of multimedia book reviews to increase elementary students' independent reading. *Reading Research Quarterly 35*(3), 384-419.

Rockman, S. (2000). A lesson from Richard Nixon: Observations about technology policy and practice in education. The Secretary's Conference on Educational Technology 2000. Retrieved from October 15, 20000 from the World Wide Web: www.ed.gov/Technology/techcof/2000/rockman_paper.html

Rogers, E. (1995). Diffusion of Innovations (4th ed.). New York: The Free Press.

Rogers, P.L. (2000). Barriers to adopting emerging technologies in education. *Journal of Educational Computing Research, 22*(4), 455-472.

Ruberg, L. (1993). The impact of digital technologies on the elementary school classroom. In Virtual Literacy in the Digital Age: Selected R*eadings from the Annual Conference of the International Visual Literacy Association*, Rochester, NY. (ERIC Document Reproduction Services No. ED370566).

Schneiderman, B. (1998). *Designing the user interface* (3rd ed.). Reading, Massachusetts: Addison Wesley Longman, Inc.

Sia, A. (1992, February). *Enhancing instruction through software infusion*. Paper presented at the American Technology Education Conference, Orlando, FL.

Sherry, L., & Billing, S.H. (2002). Redefining a "virtual community of learners". *Tech Trends,46*(1), 48-51.

Sherry, L., Billig, S., Tavalin, F., & Gibson, D. (2000). *New insights on technology adoption in communities of learners*. In: Society for Information Technology & Teacher Education International Conference: Proceedings of SITE 2000. San Diego, CA. (ERIC Document Reproduction Services No. ED 444 565)

Stecher, B., & Chun, L. (2001). *School and classroom practices during two years of education reform in Washington state* (National Center for Research on Evaluation, Standards, and Student Testing/RAND). Los Angeles: University of California.

Willis, J., Thompson, A., & Sadera, W. (1999). Research on technology and teacher education: Current status and future directions. *Educational Technology Research and Development, 47*(4), 29-45.

Wood, J.M. (2000). Innovative teachers hindered by the green-eyed monster. Retrieved from the World Wide Web March 3, 2002. http://www.edletter.org/past/issues/2000-ja/innovative.html

Rethinking Disciplinary Theories & Histories: Applying *paradigms of knowledge* to Instructional Technology

Saeid Roushanzamir
University of Georgia

"A paradigm is essential to sceintific inquiry"[i]

Abstract

The purpose of this essay is to explore the relationships between major theoretical movements and selected influential leaders within the discipline of Instructional Technology. We first confronted contradictions that exist amongst researchers, including fundamentals such as the definition, history and essential identity of IT. Next, we explored how other researchers have described the IT field. Some scholars use an historical timeline to help delineate the field; others describe paradigms of knowledge. In this paper we adapt both strategies, combined with a proposed model, to define IT with an historical specificity that is both chronological and dialectical.

Ferment in the Field of Instructional Technology

To attempt a definition of the discipline of Instructional Technology (IT) is to enter a fray conducted in a muddled quagmire. Despite the existence of numerous academic programs and professional positions in the private and public sectors, the discipline, Instructional Technology, seems to retain its status of permanent discombobulation. Important summary editions of books that provide contemporary comments on the state of the discipline differ regarding what to call the discipline, when it started, how to approach it, how to conduct research and on the relationship between research and practice.

The definition of IT authorized by the Association for Educational Communications and Technology (AECT) reads: "Instructional Technology is the theory and practice of design, development, utilization, management, and evaluation of processes and resources for learning."[ii] Despite the apparently definitive statement, experts' opinions illustrate the still existing confusions.

> We have so many choices -- constructivism or neo-orthodoxy? Qualitative or quantitative? Design or postmodern critique? I imagine the diversity's good and helps us be more resilient, but the coherence of the field suffers for it.[iii]

> I am … involved in exploring implications of semiotic theory for instructional design and performance technology, which allows me to bring together a number of separate interests. Semiotics is something of a metatheory. It provides a synthesis of qualitative and quantitative research paradigms.[iv]

> [The future of Instructional Technology can be described thusly:] Online education will become global, mobile, and multimedial. Global, because communication costs are virtually independent of distance. Mobile, since mobile technology is becoming omnipresent. And Multimedial, as a result of the continuous growth in bandwidth and processing capacity[v]

One scholar concludes:

> Although we may eventually be able to conduct valid, socially responsible analytic studies in instructional technology, that time has not yet arrived. Is instructional technology research socially responsible? At the present time, it is not. Are we asking the wrong questions? For the most part, yes. Can we change this sad state of affairs? Of course, if we have the will! ….A major benefit of systemic research in education is that it yields new questions and nurtures the development of new theory.[vi]

In short, as a perusal of the Instructional Technology Research Online (InTRO)[vii] demonstrates, even long-time scholars and practitioners are beset by the strong sense that confusion reigns in our field. Given their responses, one is hard put to even use the term discipline! Some have even written that "[p]erhaps it is time to accept that IT is a subset of other varied disciplines, and not a discipline by itself."[viii]

Definitional dilemmas are grappled with by novices in many disciplines as part of their indoctrination into their particular specialized knowledge base. One IT student's attempt to define our field began with this statement: "Instructional technology is simply technology used in instruction."[ix] My own first attempt was much less elegant.[x] Different as these definitions are, the comparison among them is suggestive because it can help introduce the main themes of this paper. Those themes are: the fundamental issues of the IT discipline, its definition, history and purpose, will benefit from the consistent application of the concept of *paradigm* as proposed by Thomas Kuhn and explicated by Egon Guba and Yvonna Lincoln, to its intellectual history; and the discipline must described with an historical specificity that is both chronological and dialectical.

The next section describes and interrogates approaches to conceptualizing the IT field; it asks whether and how a time line approach can satisfactorily delimit instructional technology as a discipline. The third section describes and critiques the intellectual history of IT focusing in particular on the use of one concept: paradigm. It argues that the concept has been used in so many different ways that its usefulness in delimiting our discipline or suggesting future directions has been muted. Drawing on the insights of Guba and Lincoln, a new model of IT is proposed. This model's dimensions include the notion of research paradigm (identified by fundamental ontological and epistemological assumptions) and a notion of history that goes beyond the progressive and descriptive. Finally in the discussion I suggest that this model can help define IT by revealing its lively and open-ended history; the model by acknowledging the relevance of fundamental ontological and epistemological assumptions may also encourage theoretically informed research. It is my belief that this exercise, i.e. creating a dynamic timeline by combining that visual tool with the new model, will also reconnect theory with practice. Used together, the timeline and model highlight the basic relationships and experiences within the discipline and between IT and the social formation within which it is practiced.

Timelines of Instructional Technology Theory & Research

Can the discipline of Instructional Technology be pictured on an historical timeline and would a timeline prove helpful in moving towards a more coherent definition of the discipline? Apparently many believe so. A Google search for "instructional technology" and "historical timeline" yielded 107 results with a number of examples of individual and classroom attempts to create such timelines (and an even greater number of irrelevant hits). The following three timelines exemplify the different approaches offered.

"Methods and Tools for Communication and Learning From Prehistory to the Present"[xi] was created at the University of Illinois as a timeline beginning "Way Back When" in which methods and tools for communicating and learning include cave art, the abacus ("the first calculator"), eyeglass lenses, and the printing press. The next section, 1600s – 1700s, introduces innovative educators (among others topics), followed by the 1800s (hearing aids, Braille, telegraph, typewriter). The 20th century is broken into early, middle and late time frames with the last inclusion being a "Future TimeLine." When the "Future Timeline" is clicked one finds the "Educational Technology Timeline;" it begins in 40,000 BC and ends at 2203.[xii] Both the first timeline with its focus on "Methods and Tools" and the "Educational Technology Timeline" (as well as other links provided within both) focus attention on creative interpretations of what constitutes instructional technology. Resources drawn upon are not limited, as in Maier's chart (discussed below), to the strictly disciplinary. For example in year 2039 the creators, borrowing a phrase William Gibson's cyberpunk science fiction, foretell a time when "jacking in" becomes the common mode of existence (i.e., joining the Global Neural Network to access or control data). The audience for this timeline is primarily students at a variety of levels.

A second example comes from the University of Indiana—Bloomington's School of Education Instructional Systems Technology program. Called "A history timeline 1913-present", this document provides a history of instructional technology within an institutional (i.e. Univ. of Indiana) framework.[xiii] It too is directed towards current and prospective students since it is part of the program's promotional package.

The third example, found on the Instructional Technology Resource web site[xiv] is maintained by

David J. Maier, an Instructor of Computer Information Systems at Henry Ford Community College, Dearborn, Michigan. His web site's audience is 3-fold: students and professionals, potential clients and the interested public. Included among other useful pages and links is one page titled "Instructional Technology Timeline," arranged by decades, beginning with pre-1910s through the present.[xv] Conceptually, although decades are enumerated, this is less a timeline than a chart within which each decade is characterized by broadly defined directions of concern. For example, the 1930s is characterized as the decade when interest focused on behavioral objectives and formative evaluation. The chart provides a snapshot of large social forces (e.g. Depression and the Progressive movement) and significant studies are cited. Following the "timeline" is a "Summary of Areas" in which very broad theories are identified and described in terms of leaders, definitions and arguments, design implications and trends. But despite its title, the chart does not share either the visual or the conceptual characteristics that usually characterize a timeline. It is a useful enumeration of the research literature and thus provides a handy guide to the labyrinthine field.

The search also located the timeline which was the first step resulting in the writing of this paper.[xvi] The intention of our historical timeline was to visually represent major theoretical movements and selected influential leaders within the discipline of Instructional Technology from the twentieth century to the present. To produce our timeline we consulted three other online efforts to create a visual representation of the field of IT. One was the third example discussed above; the other two were an "Instructional Development Timeline"[xvii] and "A hypertext history of Instructional Design."[xviii] The former begins with pre-1920s through year 2000 and is more descriptive; the latter covers the same timeframe but is more analytical.

It became apparent that in each case, including our own, timelines tend to reinforce one version, one interpretation of the field's history. Alternative interpretations of the field are suppressed. We had hoped that by characterizing the history of Instructional Technology by dominant paradigms of knowledge as well as by decade, we would be able to demonstrate links between significant shifts in theory construction, broad intellectual movements and most crucially, practices, i.e. lived experiences, in the field. The naming of distinct paradigms provided another way to conceptualize and organize those movements and practices. And it was not meant to suggest unity in the discipline but only to highlight dominant thoughts and trends.

However, timelines, as a visual representation, are like a map without an interpreter. We sensed that our TimeLine was too ordered, too neat and tidy; it obfuscated contradictions and anomalies despite our attempt to use paradigms to illustrate a dynamic history. The timeline still seemed to suggest that with the inexorable march of time a concomitantly inexorable march of intellectual history occurs with very little real human agency.

In rethinking the timeline as a useful pedagogical tool we returned to Kuhn's original insights about the ways in which types of knowledge become either legitimate or peripheral. For a timeline to adequately represent the historical breaks that a new paradigm signals, the timeline should be able to demonstrate human agency, the ways in which social actions such as coalitions of power impact the acceptance or rejection of what constitutes knowledge.

Paradigms of IT Research as Identified in Key Works

Here is a definition of 'paradigms' according to Kuhn in his *Structures of Scientific Revolutions*:

" 'normal science' means research firmly based upon one or more past scientific achievement, achievement that some particular scientific community acknowledge for a time as supplying the foundation for its further practice…. These achievements "served for a time implicitly to define the legitimate problems and methods of a research field for succeeding generations of practitioners. They were able to do so because they shared two essential characteristics. Their achievement was sufficiently unprecedented to attract and enduring group of adherents away from competing modes of scientific activity. Simultaneously, it was sufficiently open-ended to leave all sorts of problems for redefined group of practitioners to resolve.

Achievements that share these two characteristics I shall henceforth refer to as 'paradigms', a term that relates closely to 'normal science.'[xix]

According to Kuhn's *Structures of Scientific Revolution*[xx] a paradigm of knowledge dominates within a scientific discipline(s) for a period of time. Dominant paradigms structure what is accepted as legitimate knowledge determining questions asked, methods accepted and so forth. This occurs because researchers share the common and often unspoken assumptions of that paradigm. Their interpretations of discoveries are constructed through those assumptions. When new discoveries, inventions, or anomalies arise that cannot be explained or which contradict fundamental assumptions a competing paradigm that more satisfactorily explains and/or predicts phenomena emerges. This period of emergence may be considered one of unrest lasting until a coherent paradigm becomes dominant.

Kuhn's insight resists the force of established truths by explaining that each successive dominant paradigm should only be understood as containing the force of current established modes of thinking and producing knowledge. This insight has been accepted in the IT field as the following illustrates:

In mature science, one paradigm typically dominates. *Progress* occurs when this dominant paradigm, unable to account for growing number of anomalies discovered in the course of normal scientific inquiry, is replaced or "overthrown" by a competing paradigm. (italics not in originl)[xxi]

However, if it is the case that no paradigm should be taken to contain a truth-value, then the metaphor of progress when discussing paradigms is also problematic. Although it may be analytically expedient to examine paradigms across time, chronology should not imply progress in motion. As described above, the advantages of using a timeline are several Perhaps most notably, they are visual and concise; they satisfy the normative desire for timely organization. The disadvantages are at least three: over-simplification, deflection of contestation and struggle to progress and order, and because they appear to be anchored by specific dates, their ideological biases tend to remain hidden.

Hannafin and Hannafin (1995) compared the status of instructional technology research in the decade beginning in 1980 concluding that empirical research and scholarship have not ranked high on the agenda of practitioners or teachers.[xxii] They note that it has even been suggested and rather than produce original research, "IT may be better served by formalizing methods for interpreting the implications of related R & D generated elsewhere than by attempting to produce its own."[xxiii] They conclude that a "pragmatic assessment of needs, means and ends" is what the Instructional Technologist does best. That conclusion seems to suggest that wherever applied, i.e. primary school, large corporation, military or academia, pragmatism is the guiding strength of the field of instructional technology.

In contrast, Driscoll, in the chapter that follows Hannafin and Hannafin in Instructional Technology claims that "[i]nstructional systems *is* … a developing science" where numerous research paradigms are competing for dominant status. (italics in original)[xxiv] Driscoll cites Kuhn's discussion of paradigm emergence and challenge to make the case that "the field should embrace a wide variety of research paradigms and not yield to the dominance of any one."[xxv] In establishing the case for IT research Driscoll comments:

When planning research and deciding among paradigms… the question is: Which problems most urgently require solutions? It is important to answer this question because adherence to particular research paradigms may affect which problems we are ultimately able to solve.[xxvi]

Similarly, Robinson asks "[w]hat is the paradigm shift that has resulted in qualitative methods becoming more important to our field…."[xxvii] At times Robinson, like Driscoll, associates the term paradigm with research methods; she also distinguishes between positivist and naturalistic inquiry and the "nonscientific paradigm" which is linked with a wide variety of methods: linguistic analysis, phenomenology, case study, grounded theory to name only a few.

Koschmann uses the term paradigm as a tool to examine specific research on instructional technology. [xxviii] He suggests that the guiding paradigms of IT have been computer assisted instruction, intelligent tutoring system, Logo-as-Latin and computer supported collaborative learning. On the other hand, Molenda and Harris assert that IT research to be analyzed correctly must be situated within each of its primary sectors: K-12 education, higher education and corporate training. "Some generalizations are possible but only within sectors."[xxix] As they correctly observe:

The spotlight is shifting toward patterns of *use* and the consequences of those patterns of use: Precisely how are teachers employing computer-based media in the instructional process? What effect is this use having on Student performance, on the role of the teacher and professor, on compensation practices, on intellectual property rights, on power relationships. These effects are much more difficult to observe and measure [than counting numbers of computers or connections].[xxx]

Hannafin and Hannafin ask that IT researchers focus on obvious daily issues encountered within IT. Driscoll calls for a focus on problem-solving. Robinson would like the "paradigm debate [to] be declared a draw."[xxxi] Robinson writes that "[a]ll ways of knowing and all social constructs should be equally accepted and represented in our literature."[xxxii] Koschmann, by evaluating methods within IT practice hopes to create order within the field.

However, the notion of paradigm as used by Kuhn, the one that this project finds helpful, exists at a level of abstraction that these authors, perhaps for good reasons, seem to omit. "…[T]he basic belief system or world view that guides the investigator, not only in choices of method but in ontologically and epistemologically fundamental ways"[xxxiii] is a clear working definition of the concept of a paradigm.

Guba and Lincoln provide a concrete discussion of competing paradigms in which they distinguish among them on the basis of a variety of criteria. Most importantly they assert that because a paradigm is "a set of *basic beliefs* (or metaphysics) that deals with … first principles."(italics in original)[xxxiv] These basic beliefs form a researcher's worldview, one that is adhered to consciously or not.

The basic beliefs that define inquiry paradigms can be summarized by the responses given … to three fundamental questions, which are interconnected in such a way that the answer given to any one question, taken in any order constrains how the others may be answered….. [Those questions are:] The ontological question…. The epistemological question…. [and the] methodological question….[xxxv]

Basic beliefs, while they may be well and coherently argued, are in the end not subject to proof. To the contrary, they are subject only to faith because there is no way to establish their truth-value. At the level of ontology, it is one's belief about the form and nature of reality which predicates what it is possible to know. The epistemological issue, i.e. the nature of the relationship between the knower and what can be known, is constrained by ontology. "How can the inquirer… go about finding out whatever he or she believes can be known?"[xxxvi] is the methodological question which again in constrained by answers to the other two questions. Guba and Lincoln delineate four main inquiry paradigms that have guided social science research: positivism, post-positivism, critical theory and constructivism. Another analysis has suggested: positivism/post-positivism, Frankfurt School/Critical Theory, neo-Marxian, and post-modernism and a third, administrative, liberal-pluralism and critical theory.[xxxvii] In all of these attempts to name and describe dominant and emerging paradigms, the ontological issues determine what the researcher thinks he or she is doing and either legitimizes or de-legitimizes specific questions, methods, interpretations and applications.

Paradigm differences exist at a philosophical level but, as Guba and Lincoln argue, these differences have practical consequences as well. They list nine issues which help demonstrate that a researcher's paradigm leads to specific choices in practice: inquiry aim, nature of knowledge, knowledge accumulation, criteria for judging quality of the research, role of the researcher's values, whether ethics are extrinsic or intrinsic to the researcher, voice of the researcher, training, possibility for accommodation among paradigms, and acceptance of the paradigm within and without the field of inquiry.[xxxviii] The ways in which these criteria are practiced by researchers are indicative of the signposts of basic assumptions and values relative to their research and fundamental belief system.

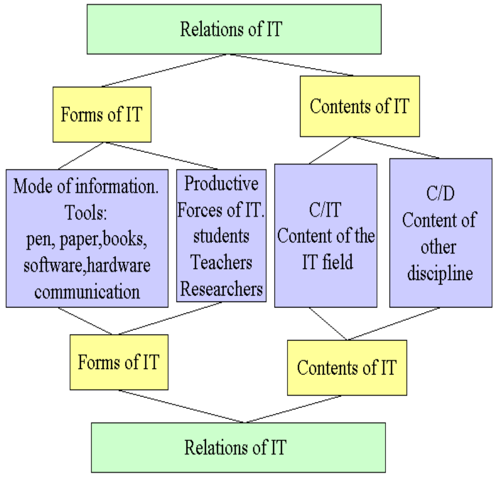
As stated in the introduction, the goal of this project is to create a vision of the discipline of Instructional Technology that is historically valid and that also illustrates the dynamic interactions of IT theory and practice. The timeline approach is clear and concise. But it tends to flatten history; its linearity tends to erase struggle and contestation. In the following section a new model of the discipline is proposed. I hope the model enhances the chronological visual tool, i.e. the timeline, it encompasses and expresses intra- and extra-paradigm critiques.

Proposed Model of Instructional Technology

Instructional Technology can be comprehended as a **web of forms** that delivers a **web of contents**. The interactions between the two produce and reproduce what I call **relations of Instructional Technology**. Each of these concepts is explained in the following paragraphs. The Forms, Contents and Relations produce knowledge for the subject: us, i.e. researchers, practitioners and all those who encounter and use instructional technologies. Whose knowledge is produced and reproduced? Fundamentally, that knowledge is one part of the larger social formation, i.e. the reality of concrete historical social and economic conditions. A definitional model using the three concepts, i.e. Forms, Contents and Relations, is important because it will help historicize the discipline of Instructional Technology. This in turn helps researchers and practitioners analyze how and in whose interests instructional technologies are developed and used.

Media, across the gambit of chalk to pen and paper to computers and software no mater how simple or how sophisticated, can be categorized Forms for education. The tools, whatever they are, are the mode of information of the Forms. Forms of IT also include messengers and message takers; these are the productive forces of Forms. It is the responsibility of IT as a field to understand these means and their functions. Thus, Forms includes the mode of information and its productive forces.

Figure 1.



Contents is a two-pronged concept that includes the content of the field of IT (**C/IT**) and the content of other disciplines (**C/D**). **C/IT** as content in and of itself is a relatively new discipline, but nevertheless, a discipline since it is characterized by an identifiable body of theory and research (although with contradictory research and underdeveloped theory). **C/D** signifies the content of all disciplines. IT is not responsible for the contents of other disciplines; nevertheless, IT has been delivering the forms to those contents. So in this sense Instructional Technology can be conceptualized as a long-established field. The current ferment in the newly emerging discipline is the result of attempts to discover and articulate its boundaries systematically.

Relations of IT to its Forms and Contents (**C/IT** & **C/D**) are determined by the ownership of the mode of information of Forms and distribution of Forms within any particular historical moment. Furthermore, Relations of IT is also determined by the ownership of Contents since **C/IT** and **C/D** are also historically specific to a society.

This definitional model (see figure1) is informed by a structure of binaries and by a commitment to an historical perspective. The difference between a field of inquiry and a scholarly discipline is located by the distinction between individual and structural variables. Anyone, scholar or armchair philosopher, can define their own field of inquiry. However, a discipline exists only as a series of relationships; those relationships encompass individuals, but also group and institutional aspects. Within Instructional Technology, relations are determined by two sets of characteristics: first, the ownership and distribution of the mode of Forms within a particular historical moment and second, the ownership of IT's content.

The following set of binary oppositions will help move this model forward. On one side, consider the terms: critical thinking, educated, process oriented, and scientist. On the other side consider the terms: functionalist thinking, trained, task oriented, engineer. Together, these binaries form structures of experience and meaning with particular relevance to some of the ongoing (seemingly never ending!) debates within our field.

One way to imagine their relevance is to locate the Forms, Contents and Relations of IT. When the largest supporters and developers of IT are located in corporations and the military we can claim with some assurance that the knowledge being produced and reproduced falls on the functionalist side. Thus, the advantage of insisting on historically particular theory and research is a recognition that the future in both unknown and malleable. There is no absolute progress; there are always opportunities for reform or radical change. One cannot deny the power of entrenched ideas and practices; but a concrete historical analysis also always shows that there are multiple possible futures.

Figure 1. Definitional Model for Instructional Technology

Paradigms of IT Knowledge: Historicizing Theory

Armed with the ideas previously discussed, it can now be argued that the field of Instructional Technology has experienced three distinctive inquiry paradigms from early 20th century to the present: Positivism/Post Positivism (1900-1930), Administrative (1940-1970), and Constructivism (1980-Present).

Positivism/Post Positivism Paradigm In IT (1900-1930)

During this period, the pioneering work of a few major scholars such as E.L. Throndike, John Dewey, Maria Montessori, Franklin Bobbitt, and Ralph Tyler directed and focused the burgeoning fields of social science, including IT. Despite differences in their approaches to education, these scholars nonetheless shared a common belief structure. They were objectivists applying methodologies in which verification of given hypotheses was sought using quantitative and qualitative methods. They also shared the view that individual betterment was a necessary condition for social betterment. In the following table this paradigm is described in terms of chronology, characteristics and known influencers.

Date	Characteristic	Known influencers
Pre 1920s	Emergence of Social Science, as both legitimate & legitimator	E.L. Throndike John B. Watson John Dewey Maria Montessori
1920s	Individualism Atomization	Franklin Bobbitt Sidney Pressy Fredric Burk Carlton W. Washburne Helen Parkhurst
1930s	Emergence of Social Effects	Ralph Tyler Alan Turing C.F. Hoban F. Dean McClusky

Administrative Paradigm in IT (1940-1970):

By that late 1930s the emphasis on individual betterment as a direct path to a better society succumbed to organizational and institutional thinking. The term “administrative paradigm” is used to signify a break with naïve positivism together with a top-down approach to thinking about social issues. It is also during this period that Instructional Technology emerged as a distinct field.

In this period some of the influential leaders and researchers were still identified with established social sciences, but some identified themselves as IT scholars. Scholars for this era include B.F. Skinner, Edgar Dale, Benjamin Bloom, Jean Piaget, Robert Glaser, Robert Gagne M.D. Merrill to name a few.

The shift from positivism to a committed post-positivism also signaled the shift to organizational and institutional themes. These leading theories such as behaviorism and cognitivism held sway in their respective disciplines but only as models to be tested. However, IT researchers were more likely to simply lift findings to graft onto their own work as a pragmatic expedient. In the case of instructional technology, at least during this period, it was large social imperatives rather than individual betterment that drove the questions asked. This can be explained in a number of ways. Perhaps first among those explanations: instructional technology’s immediate use value to the military efforts coincided with the emergence of IT scholarship and its institutionalization in higher education. Only by the latter years of this period were the questions that posed a significant challenge to administrative research, questions that had been raised and seriously considered in other fields, asked with IT. The following table shows subclasses of the administrative research paradigm.

Date	Characteristic	Known influencers
1940s	Administrative & Pragmatic	Edgar Dale Kurt Lewin
1959s	Behaviorist & Cognitive	B.F. Skiner Benjamin Bloom James Finn Jean Piaget John Flanagan
1960s	System Approach	Robert Glaser Robert Gagne Jerome Bruner Norman Crowder Donald Ely M.D. Merrill
1970s	Challenges to System Approach	David Ausubel Seymour Paper

Constructivist Paradigm in IT (1980-Present):

Without being called a technological determinist, it seems fair to argue that the rapid proliferation of ever-smaller and increasingly sophisticated new technologies helped thrust the unexplained anomalies of administrative research to the forefront of the IT field. Administrative research was increasingly unable to provide satisfactory answers for questions that faced IT researchers (e.g., questions regarding the mechanism of learning processes, etc.) and constructivism seemed to promise a way of beginning to find more satisfying answers by posing better questions. The following table shows a selection of the influential scholars and characteristics of the past two decades in our field.

Date	Characteristic	Known influencers
1980s	Competing Paradigm	Brent Wilson David Kolb Michael J. Hannafin David H. Jonassen Robert Heinich
1990s	Constructivism	Bernie Dodge Gloria Gery Walter Dick

Conclusion: Inquiry Paradigms & the Definitional Model

The goal of this project is to tell the history of a discipline, Instructional Technology, through a medium that is visually clear and historically dynamic. The original timeline had adapted Kuhn's concept of paradigm into a decade-by-decade conventional timeline. However, this approach seemed at once to create overly abstract categories and overly simplistic distinctions. The design of a new definitional model helps animates the timeline because the model itself provides a structural frame; in short, it acknowledges that all inquiry operates within categories of the dominant power structure. Each paradigm is distinct because to work within each demands an ontological commitment that the researcher may or may not be able to recognize or articulate. An historical timeline is itself an ideological tool, i.e. to comprehend the form requires acceptance of a certain kind of knowledge. Certainly, once recognized its ideology must be dealt with but it would be foolish to put the tool aside; ideology is a reality of all social relations, e.g. economic, political, social and cultural and at every level of social interaction, e.g. institutional, organization, and individual. The Definitional Model recognizes the social formation within which the field of instructional technology exists; it inserts real relations of experience, i.e. the concrete power differentials that are produced by economic capital (ownership of forms) and cultural capital (productive forces of forms). It helps identify instructional technology's subject matter by recognizing that the interactions of and between **C/D** and **C/IT** must each be taken into account rather than reified as a false distinction or overlooked altogether.

All of these aspects interact as a series of constantly changing interactions; the relations of IT both interact with and mimic the ongoing negotiations of social and political relations within the social formation. The intellectual history of our discipline, no less than any other history, involves a process of recovering and reclaiming pieces of human thoughts and actions. Such a history should be as open as possible to the nuances of individual agency without underestimating structural determinants.

----

[i] Pajares, Frank. (date) "Outline and Study Guide of The Structure of Scientific Revolutions by Thomas Kuhn," at website: http://www.emory.edu/EDUCATION/mfp/Kuhn.html

[ii] Association for Educational Communications and Technology web site. http://www.aect.org/standards/knowledgebse.html

[iii] Brent Wilson, 1997 interview; webpage of Instructional Technology Research Online (InTRO), http://www.gsu.edu/`wwwitr/

[iv] Marcy P. Driscoll in an interview, n.d.: ; webpage of Instructional Technology Research Online (InTRO), http://www.gsu.edu/`wwwitr/U

[v] Morten Flate Paulsen, 2003 interview; webpage of Instructional Technology Research Online (InTRO), http://www.gsu.edu/`wwwitr/

[vi] Thomas C. Reeves, "Questioning the Questions of Instructional Technology Research," [Online] Available http://www.hbg.psu.edu/bsed/intro/docs/dean/ February 15, 1995.

[vii] This web site is maintained at Georgia State University; Editorial responsibilities for InTRO reside with Steve Harmon, marshall Jones, and Dan Surry who describe the site thusly: "InTRO is dedicated to providing professionals in the field of Instructional Technology with an electronic forum to disseminate, discuss, and advance research in Instructional Technology and related fields. InTRO takes advantage of the versatility of electronic communications to widely disseminate a broad range of works in a variety of media through fast, convenient, and accessible methods. This communication is intended to keep professionals in Instrucitonal Technology and related fiels informed, while providing researchers a forum in which they may receive critical and constructive feedback on their work." http://www.gsu.edu/`wwwitr/

[viii] Michael J. Hannafin and Kathleen McDermott Hannafin, "The status and future of research in Instructional Design and Technology Revisited," in Instructional Technology: Past, Present and Future, 2d. ed. Edited by Gary J. Anglin, Engleweek, CO: Libraries Unlimited, 1995, p. 320.

[ix] M. Miller's web site. http://www.arches.uga.edu/`mlmiller/definition/index.html (this was part of an assignment in Univ. of GA EDIT6100 class, 2001)

[x] S. Roushanzamir's web site. http:/www.arches.uga.edu/`roushan/EDIT6100/start.html (this was part of an assignment in Univ. ofGA EDIT6100 class, 2002)

[xi] 08/99, wads copyright 1999, Web Technology Group, Universtiy of Illinois,

[xii] http://cter.ed.uiuc.edu/cter2/ci335/timeline.html

[xiii] http://www.indiana.edu/~ist/students/history/timeline.html

[xiv] On the web site of David J. Maier, http://www.dmaier.net/index3html (for introduction) or go to: http://www.ittheory.com/index.html

[xv] http://www.ittheory.com/timelin2.htm

[xvi] On the web site of Saeid Roushanzamir, "Instructional Technology Paradigm Timeline," Saeid Roushanzamir and Emily Giles. http://www.arches.uga.edu/`roushan/timeline/

[xvii] From the website of Leilani Carbonell, http://www.my-ecoach.com/idtimeline/

[xviii] http://www.coe.uh.edu/courses/cuin6373/idhistory/index.html

[xix] Thomas S. Kuhn, The Structure of Scientific Revolutions, 2d ed. Chicago, U Chicago, c1970.

[xx] Thomas S. Kuhn, the Structure of Sceintific Revolutions, 2d ed. Chicago, U Chicago, c1970. See also 3rd ed.k and Frank Pajores, "Outline and Study Guide of *The Structure of Scientific revolutions* by Thomas Kuhn," at http://www.emory.edu/EDUCATION/mfp/Kuhn.html

[xxi] Mary P. Driscoll, "Paradigms for research in Instructional Systems," In *Instructionals Technology: Past, Present, and Future.* Ed. by Gary J. Anglin, Engleweed, CO: Libraries Unlimited, 1991.

[xxii] Michael J. Hannafin and kathleeen McDermott Hannafin, "The status and future of research in Instructional Design and Technology Revisited," in INstructional Technology: Past, Present and Future, 2d. ed. Edited by Gary J. Anglin. Engleweed, CO: Libraries Unlimiited, 1995, 314-21.

[xxiii] Ibid. p. 320

[xxiv] Marcy P. Driscoll, "Paradigms for Research in Instrctional Systems," in Instructional Technology:

Past, Present and Future, 2d. ed. Edited by Gary J. Anglin. Englewood, CO: 1995, 322-29.
[xxv] Ibid. p. 322
[xxvi] Ibid. p. 332
[xxvii] Rhonda S. Robinson, "Qualitative Research-a case for case studies," , " in Instructional Technology: Past, Present and Future, 2d. Edited by Gary J. Anglin, Englewood, CO: Libraries Unlimited, 1995, 330-39.
[xxviii] Timothy Koschmann, "Revisiting the paradigms of instructional technology," in Meeting at the Crossroads: proceedings of the annual conference of the Australian Society for Computers in Lerarning in Tertiary Edcuation, 2001.
[xxix] Michael Molenda and Philip Harris, "Issues and Trends in Educational Technology," in *Educational Media and Technology Yearbook* Volume 26 edited by robert Mribe Branch and Mary Ann Fitzgerald. 2001. p. 3-15.
[xxx] Ibid. p. 12
[xxxi] Ibid. p. 332
[xxxii] Ibid. p. 332
[xxxiii] Egon G. Guba and Yvonna S. Lincoln, "Chapter 6; competing paradigms in Qualitative Research" in *The Landscape of Qualitative Research,* edited by Norm Denzin and Yvonna S. Lincoln.
[xxxiv] Ibib.
[xxxv] Ibid. p. 108
[xxxvi] Ibid. p. 108
[xxxvii] Stuart Hall; Todd Gitlin
[xxxviii] See Table 6.2 "Paradigm Positions on Selected Practical Issues" in Guba and Lincoln, ibid., p. 112

From Socrates to Cyberspace: Enhancing Teaching and Learning through Technology

Heidi L. Schnackenberg
Margaret D. Maughan
Elizabeth Zadoo
Plattsburgh State University

Abstract

This initiative attempted to promote university faculties' use of technology in their teaching practice by equipping them with the latest techniques and instructional strategies. Faculty participants were given an initial workshop with a noted educational technology expert and then subsequently paired with a student partner with whom they were to develop a technology-enhanced lesson plan. At the end of two semesters, participating faculty implemented lessons that integrated technology ranging from utilizing a course website to engaging in online discussion forums. The most intriguing implication of this research was that although only four faculty members were able to complete a lesson, over 100 students were reached among the various classes taught by these professors. When contemplating faculty professional development, it is most important to remember that even if only some faculty members engage in these programs, many students can benefit from the knowledge and skills that their professors acquire.

Overview

Extensive research in the field of college/university faculty professional development has been done over the past decade (Roberts & Associates, 1999; Mihkelson, 1997; Cohen, 1995; Valadez & Duran, 1991). Specific findings suggested that a challenge existed to the evolvement of college-wide technology standards and the expansion of learning resources for technological collaboration (Gillan & McFerrin, 2001). The empowerment of personnel through partnership was highlighted in three New York studies (SUNY Series "Frontiers of Education", Petrie, Ed., 1995; the MTIP Study, Ware, 1992; Niagara County Community College, Harnish, 1986) as well as others outside the state (Bell, 1996; Luna & Cullen, 1995). The role of professional development was investigated as a significant element in the promotion of change (Gaston & Jackson, 1998; Guskey & Huberman, 1995; Guskey, 1994). Additionally, faculty vitality and its relationship to professional development have been studied (Brown, 1996; Kelly, 1990).

The State University of New York (SUNY) directive that mandates accreditation of teacher education programs for all campuses has necessitated the in-depth examination of the all teacher education programs and education faculty professional development. The process of accreditation requires comprehensive curricular reviews to determine the extent of compliance with the individual national associations (National Council of Social Studies, National Council of Teachers of English, National Society for Teachers of Mathematics, National Science Teachers Association, for example). Each organization has a series of standards that define the goals and objectives for teachers and teacher candidates. Examination of faculty professional development is also part of the process. This program targeted faculty integration of technology into pedagogical practices.

While this type of professional development activity has occurred previously (Schwier et al, 2000; McMullen, Goldbaum, Wolfe, & Sattler, 1998; Pierson et al, 2001) past efforts have fallen short primarily in the area of continued support (Strudler & Wetzel, 1999) both with the utilization of technology and the implementation of pedagogical techniques (McAlister & Reagan, 2001; Burkett, Macy, White, & Feyden, 2001; Smith & Kolosick, 1996; Baldwin, 1999). The current program attempted to fill a void in the professional development activities existing in the field of higher education by addressing the characteristics of adult learners through the incorporation of a unique mentoring relationship between faculty and undergraduate teacher education candidates.

Current research in the area of faculty-student mentoring defined the process as "a creative alternative to direct instruction and teaching that provides an emotionally supportive relationship" (Shaughnessy, 1995). Studies indicated that a positive interaction within the school culture (Pitton, 2000; Portes, Longwell-Grice, & Chan, 2000; Mullen & Lick, 1999; Tauer, 1996) produced tremendous benefits to the participants. The use of mentoring within the university setting was also studied (Gaffney, 1995; Henderson & Welch, 1993).

While faculty members are certainly experts in their content areas, as well as pedagogy, undergraduate candidates participating in this project were identified as highly technology literate and therefore had some degree of expertise themselves. By forming faculty-candidate mentor pairs, faculty were able to feel that they had support in the development of their projects at the same time that candidates were able to reveal to faculty the teaching and/or learning options that various technologies can offer. Together, through the symbiotic combination of their various areas of expertise, the mentor pairs were able to work to create and implement ways to appropriately integrate technology into the university classroom. The overall impact of this program on the target group was multi-tiered. Primarily, university faculty gains facility, ability, and creativity when utilizing technology in the courses they teach. Through the effective use of technological applications in classes, these faculty members then serve as models of "best practice" for both the teacher education candidates with whom they are working as well as other students in their courses. As undergraduate candidates observe and interact with faculty in designing their lessons and fully utilizing technological capabilities, they will then incorporate these abilities into their own scholarly work and future experiences in classrooms. Potentially, as with most educational institutions, other faculty may become interested in the use of technology in pedagogy through the examples of their peers and positive reactions from teacher education candidates. Overall, an institutional culture will begin to be created where the integration of technology into teaching practices is effectively used, consistently modeled, and ultimately seamlessly interwoven into courses in content specialty, curriculum, and a variety of other areas.

Method

Participants

Program participants consisted of six tenured and two non-tenured university faculty members in education at two SUNY institutions located in northern New York. By the end of the program, two participants had dropped out, and one was not able to implement their project with a class, but did continue participating. Each faculty member was paired with a teacher education candidate who expressed an interest in the project and who demonstrated proficiency with a variety of hardware and software.

Additionally, an educational technologist served as project facilitator at each participating campus. They coordinated the technical needs with participants via their home campus media services. The project coordinators were also available on an as-needed basis for technical or pedagogical assistance.
Finally, implementation of an initial workshop was facilitated by a nationally recognized expert in the integration of technology into pedagogical practices.

Materials

The participating faculty provided the curricular materials needed for this program. As faculty members have many materials and lessons previously developed, they have the expertise and appropriate content for their specific courses. This project did not attempt to redevelop entire courses, however, it did enhance the learning and teaching experience through the incorporation of technology. Therefore, hardware and software were provided to each faculty member by participating campuses. The purchase of additional materials (electronic or otherwise) was not necessary.
Presentations given during the initial workshop were researched and created by the educational technology expert. They were comprised of original information and formed the basis for discussions about technology/pedagogy infusion, outcomes, and assessment.

Procedures

Prior to beginning the project, an Application for the Use of Human Participants, including all investigator's vitae and research instruments, was filed with the primary institution's Internal Review Board (IRB). Permission from the IRB to initiate this research was given soon thereafter.

Email and hardcopy letters were sent out to all education faculty members at participating campuses announcing and describing the professional development project. The letter detailed program expectations of all participants, and requested that individuals volunteer to participate. Information was provided as to whom to contact if interested in the program. Also, a Faculty Expectation Survey on Technology Guidelines (Schrum & Skeele, 2000) was administered via email. Hardcopy or electronic copies were returned to the project facilitators prior to the start of the program.

An initial workshop was scheduled during which a welcoming keynote speech was to be given by the educational technology expert via teleconference. As well, a "technarium" was organized where faculty could get hands-on training with various types of hardware and software (digital camera, PowerPoint, etc.). A dinner was also included. However, despite initial interest, no participants actually signed up for this workshop and it was subsequently cancelled.

A second workshop was organized approximately six weeks later and all participants attended. During this session, a noted expert in educational technology facilitated presentations and discussions surrounding the integration of technology and pedagogy, as well as outcomes and assessment. This was an all-day session and ended with faculty brainstorming ideas about how to infuse technology into a lesson. Both breakfast and lunch were provided.

Following the second workshop, faculty members were paired with a technology-literate teacher education candidate. The candidates were chosen by the campus facilitators based a demonstrated proficiency with the following hardware and software: Microsoft Word, Microsoft Excel, Microsoft Word, zip disk, burning a CD-RW, projection equipment, scanner, digital camera, and a camcorder. As well, the ability to design web pages and have facility with the Internet and various browsers was considered. These teacher candidates earned a one-semester, one-credit, course credit for working with their faculty partners for one of the two semesters over which the initiative occurred.

Subsequent to choosing qualified teacher education candidates and confirming faculty interest, mentor partnerships were formed. Candidates and faculty members were paired according to content area concentration, grade level specialty (elementary, secondary, special education) where relevant, and previous opportunities to work together (i.e. teaching assistants, research assistants). After pairing faculty and teacher education candidates, the pairs were left to work independently to formulate lesson ideas and work with the necessary technologies. Campus facilitators exchanged emails and names with the pairs and faculty were to make initial contact with their partner. Some pairs met only once while others met several times. Campus facilitators were available at all times to help both the faculty and the candidates create ideas and work with the necessary technologies. After one semester, the partnerships dissolved as many candidates were leaving campus to participate in their student teaching and field experiences.

Faculty were not given new partners during the second semester of the program as it was deemed too difficult to re-establish mentor pair relationships. Additionally, many faculty members informally expressed to one campus facilitator that it would be more expedient for them to progress with their project ideas alone and simply request assistance and support from the facilitators when needed. During the second semester of the program, faculty proceeded to complete development and implementation of their lessons in both undergraduate and graduate classes. A final meeting of participants in order to share and discuss the program was originally planned, but due to the ending of the semester it never materialized.

Criterion Measures

Faculty proficiency with a variety of technology skills was measured through a Faculty Expectations Survey on Technology Guidelines (Schrum & Skeele, 2000). These responses were used to gear the workshop subsequently given to the faculty. This survey consists of four sections (Foundations, Instruction, Management, and Ethics) with a total of six questions, each with a subset of Likert-type items. The total number of questions is 37. A sample item from the Faculty Expectations Survey on Technology Guidelines is below:

II. INSTRUCTION Each faculty member should be able to:	3c. Use strategies appropriate to instruction by understanding the variables of search engines and knowing how to select ones that are appropriate to subject.	Need to Learn	Progressing	Proficient

A 10-item item Workshop Evaluation Survey assessed faculty's satisfaction with various aspects such as materials, content, presenters, facilities, etc. of the one-day workshop in which they participated. The survey consisted of two selected-response items and eight constructed-response items.

Halfway through the initiative, faculty responded to an 11-item constructed-response Midpoint Survey. This instrument was used to assess faculty progress on the project as well as ways in which the campus facilitators could assist the faculty with their on-going projects. Sample items from the Midpoint Survey are below:

Please reflect upon and write about the following:

- State the challenges you have encountered up to this point in developing your technology-infused lesson.
- What type of help or assistance do you need at this point to make lesson implementation happen?

After faculty implemented their technology-integrated lesson, students in their classes were asked to respond to a 15-item Survey for Teacher Education Candidates. This instrument consisted of eight yes/no and seven constructed-response items. It attained student satisfaction with and recommendations for the lesson in which they participated. Each survey was administered with a cover letter attached to the front explaining that students were giving "informed consent" by taking and returning the anonymous surveys. An item sequence from the Survey for Teacher Education Candidates is presented below:

Please reflect upon and write about the following:

7. Did you feel that the technology helped you learn things that you otherwise would not have learned? Yes No
8. Why or why not?

Additionally, upon completion of the initiative, faculty completed a 10-item constructed-response Final Faculty Survey. This questionnaire assessed faculty's challenges and successes with the technology-infused lesson that they implemented. It also gathered their suggestions for future initiatives of this nature. Sample items from this survey is listed below:

Please reflect upon and write about the following:

3. Please state the challenges you have encountered in developing and implementing your teaching activity or student activity.
4. What are your suggestions or recommendations for reducing these challenges in the future?

Data Analysis

Simple mean scores for the Faculty Expectations Survey-Technology Guidelines were tabulated on a one to three point scale (Need to Learn = 1point, Progressing = 2 points, Proficient = 3 points). Responses to the Workshop Evaluation Survey, Midpoint Survey, Survey for Teacher Education Candidates, and Final Faculty Survey were tallied (respectively) and categories of responses were created. These categories were then collapsed in order to attain informative themes.

Focus groups for both the faculty and their classes were planned in order to emphasize reflection on the technology-enhanced lessons that were taught and to clarify or expand upon answers given in the surveys that warranted further investigation. However, once again due to the program ending at the end of the second term this was not possible as many faculty and most students were not on campus.

Data were not collected from teacher candidate mentor-partners due to these individuals leaving campus and difficultly contacting them.

Results and Implications

Faculty Expectations Survey – Technology Guidelines

Under the skill category I.Foundations, faculty rated themselves most proficient (M = 2.63) on items 1d. Each faculty member should be able to demonstrate effective uses of computers and technology by using available help resources (manuals, built in help, media and tech specialists, other colleagues, Internet, and students) and 2a. Each faculty member should be able to identify and select software and other non-hardware resources by locating resources using the Internet, catalogues, books, professional journals, and colleagues. While this rating does not demonstrate faculty's ability to utilize various technologies, it does speak to their resourcefulness and ability to seek out ways of problem-solving. This could be due to the honing of these types of skills in their academic preparation as well as teaching experiences. Because problem-solving skills are at the heart of learning technology skills, the advanced ability of these faculty members with this capacity makes them excellent candidates for becoming tech-savvy.

Faculty rated themselves least proficient (M = 1.5) in this same category on items 1a. Each faculty member should be able to demonstrate effective uses of computers and technology by use of correct terminology related to technology (e.g. do you know what LAN, LCD, Ethernet, or html refer to?), 1e. Each faculty member should be able to demonstrate effective uses of computers and technology by use of projection devices, set up computer for projection, connect video and sound amplification, use of wall mounted controls, use remote control, etc., and 2d. Each faculty member should be able to identify and select software and other non-hardware resources by selecting software that is compatible with institutional supported hardware platforms and software specifications. Responses to this item indicate that faculty did not feel proficient with the technology itself, compatible software or the related terminology. This makes sense in the context of this project as the three items are clearly related to one another. As well, if faculty felt confident in these areas it was less likely that they would participate in a project of this nature.

In the skill category II. Instruction, scores were highest, indicating that faculty felt most proficient (M = 2.5) with items 3c. Each faculty member should be able to use strategies appropriate to instruction by understanding the variables of search engines and knowing how to select ones that are appropriate to subject and 3e. Each faculty member should be able to use strategies appropriate to instruction by communicating and collaborating (via Internet, intranet, email, and listservs). These responses were unsurprising as utilizing search engines and communicating via email are integral parts of the job of a faculty member in academia today. It is necessary to research topics, literature, and current publications in specialty areas using the Internet. Additionally, it is virtually impossible to stay in communication with today's undergraduate and/or graduate students without utilizing email. It was encouraging that the faculty who participated in this project were comfortable with these techniques.

Conversely, faculty in this skill category felt least proficient (M = 1.1.3) with item 4d. Each faculty member should be able to develop and present lessons so that students can publish (desktop publishing, web pages). This response was also unsurprising as faculty and teachers generally utilize teaching techniques in their classrooms with which they are comfortable. Faculty who participated in this project did not feel completely proficient or comfortable with technology themselves so it was unlikely they would incorporate techniques such as multimedia into their teaching prior to completing the initiative.

Interestingly, in skill category III. Management, faculty responded most positively (M= 2.88) and felt most proficient with item 5c. Each faculty member should be able to use technology for administrative tasks by using the school's communication system (email and community announcements). However, they responded that they felt least proficient (M = 1.14) with item 5d. Each faculty member should be able to use technology for administrative tasks by using Lotus Notes discussion groups. As both items have to do with electronic communication, the disparity in the confidence of faculty with each forum is curious. Perhaps one explanation lies in the fact that faculty confident with item 3e, having to do with using email for instruction so it stands to reason that they would feel confident using it for seemingly simpler administrative tasks. As well, the skills needed to successfully moderate and/or create online discussion groups are quite advanced (in Lotus Notes or otherwise) and faculty may not have had instruction in or experience with this medium. In some ways it is actually rather encouraging that faculty were aware of the variety of skills needed for online discussion and therefore realized that it can be a complex task for which they are not ready prior to receiving specialized training. The low response to item 5d could also be in part due to the fact that some faculty may have felt proficient with discussion groups, just not in the Lotus Notes format.

The final skill category on this instrument, IV. Ethics, received the highest overall ratings of any category (M = 2.81). While overall faculty felt quite proficient in this area, they felt most proficient (M = 3.0) with item 7c. Each faculty member should be able to adhere to copyright and site license requirements by following university policy regarding the use of the server for software. They also felt least proficient (M = 2.75) with item 7b. Each faculty member should be able to adhere to copyright and site license requirements by respecting intellectual property rights (cite sources and references using proper bibliographic formatting. Though both items were rated quite highly, faculty may have felt more confident with respecting university site license policies than with utilizing proper citations formats as the former may seem more concrete than the latter. University media services are generally quite clear about the restrictions on software and hardware and they generally manage it for staff and faculty at the institution. However, it was not until relatively recently that appropriate citation formats were created and documented for electronic media. (It wasn't until the 4th edition of the American Psychological Association's Style Manual that these citation formats appeared.) For some professors, these types of citations can still be quite confusing, although the importance for giving credit for others works (i.e. citing) appears to be a priority.

Faculty Workshop Evaluation

Overall faculty response to the workshop was quite positive. Most (8) faculty felt that they received information in a timely manner and that the information they received from a colleague (7) enabled them to easily register (6).

The content of the workshop - technology/pedagogy infusion, outcomes, and assessment – was also well regarded. Four faculty members found the information very helpful and useful, five stated that the presenter was excellent, and nine said that the session met or surpassed their expectations. This feedback verified that knowledge assessment of adult learners prior to implementation of training was beneficial. Positive responses support the fact that the workshop was designed in response to participant statements given on the Teacher Expectations Survey – Technology Guidelines.

A few individuals offered suggestions that are certainly worth considering as well. One person stated that s/he would have liked more time for interacting on the web. Another said that s/he would like more emphasis on hardware and software use. Additionally, two final comments reflected that more time was needed in order to develop philosophies regarding technology and that in the future this group would like to hear speakers on technology integration changing and/or informing school culture. In subsequent faculty professional development initiatives, each of these ideas will be taken into consideration.

Faculty Midpoint Survey

The Faculty Midpoint Survey was the first opportunity in which faculty members had the chance to document their technology-infused lessons. Most of the participants had defined or begun working on a project, with only one person not having developed one. The ideas for the lessons were quite varied and interesting. One faculty member reported recording digital video images of teacher education candidates and their students (with permission of course) with the intention of using them as instructional examples in future classes. Another person stated that s/he was collecting multimedia resources and importing them into a PowerPoint presentation which explored the human mind and media. A different faculty member also reported working on a PowerPoint presentation in order to transfer his/her lecture slides into a multimedia format. Yet another individual was attempting to have teacher candidates record digital still images of museum/research field trips in order to develop subject specific units. Working on websites in order to develop teacher education candidate professional portfolios was also cited. A final faculty member stated that s/he was going to incorporate interactive grammar lessons into his/her webpage as reinforcement for concepts learned in class. Overall, the wide variety and ambition of these lessons was evidence of the willingness of faculty members to invest their time in technology. It also spoke to their creativity in planning lessons.

Faculty members also reported working on their faculty websites (3) and PowerPoint (2) with their teacher candidate partners at this juncture in the project. Three also reported that it was helpful to work with their student partner (or anyone with superior knowledge and skill) in order to learn more. It is interesting to see that these two forms of multimedia were most popular among the participants. It would stand to reason that both the preservice teachers and the university faculty saw immediate value in these forms of technology for instruction and learning. Perhaps it is the ability of both websites and PowerPoint presentations to efficiently and effectively deliver information that made them most prominent in the participants minds. While it is exciting that of all of the possible technologies from which to choose the candidates and teachers choose a multimedia format, it would be disappointing to consider that they did so because these formats could effectively take the place of or supplement lectures rather than using them for their interactive instructional properties.

When reporting the challenges the faculty participants faced midway through this initiative, the one most commonly cited was time considerations for development (5) as well as for implementation (2). This statement is supported by Sprague, Kopfman, and Dorsey (1998) who posit that effective faculty professional development initiatives for technology should include, among other things, an opportunity to explore technology integration as well as time to learn the technology itself. Faculty from this project also offered solutions to this challenge by stating that forward planning to allot needed time for gaining initial skills with the equipment would help to alleviate this problem (1). As well, one faculty member also recommended more staff development with educational technology would also be a possible way to spend much needed professional development time. While the campus facilitators were aware that faculty would need course release time to successfully create and implement their technology-infused lessons, none was granted by the administration. In an attempt to compensate for this, the facilitators asked that faculty only create and implement one lesson over the course of two semesters. However, that timeframe still appeared not be insufficient.

When asked about the positive aspects of the project, two faculty members stated that the support and guidance offered by the campus facilitators were important. Additionally, other faculty reported that they were learning new things (1), gaining insight into instructional technologies (1), thinking about new ways to use computers in their classrooms (1). These insights are especially important because they reveal that faculty not only began thinking about technology for teaching and learning, but they also revealed a sense of confidence by feeling that the campus facilitators were unfailing there to support them. This is quite key with professional development initiatives as ongoing, just-in-time support seems to be critical for successful faculty training with technology.

Finally, valuable insights were offered on ways to improve this project. While two faculty members stated that it was great as it was, others stated opportunities for intensive instruction (1) and a longer timeline (1) would be very beneficial. The fact that the suggestions for improvement again dealt with the issue of time is more support for the idea that faculty need course release and continuous support for professional development involving technology.

Faculty Final Survey

At the conclusion of the initiative, five of the participating faculty completed the Faculty Final Survey. Surprisingly, all five stated that they utilized technology in both their teaching and in student activities. Although unspoken, it was the researchers' hope that participating faculty would be able to infuse technology into either their teaching or student learning, but that perhaps doing both would be too demanding. The fact that these five individuals utilized technology in both types of activities was testimony to their hard work, motivation, and a certain level of success of the professional development program.

While faculty reported the ways in which they were utilizing technology in their classes at the time of the Faculty Midpoint Survey, some of these activities changed by the time the Faculty Final Survey was administered. Some teaching activities were reported: two individuals stated they used PowerPoint to prepare slides, while others learned FrontPage or other web editors to create websites or collate resources. Overwhelmingly, five faculty members reported using web resources (discussions, information resources) in their student activities. One also stated that s/he had the students design a program portfolio web page for outcomes-based assessment. These results are particularly important because they imply that faculty members clearly see the World Wide Web as a useful resource in education. It is the first type of technology that they choose to use so perhaps they viewed it as the most easy and/or interesting for them to learn. As well, they can recognize the value of this resource for student learning more clearly than say graphic/photo editing software or spreadsheets and databases. Also, it appears that PowerPoint is viewed more as a tool to facilitate teaching rather than student learning. Given these results, perhaps it would benefit educational technologists doing professional development to begin with the uses of the World Wide Web and Internet as it is easiest for faculty to see the value of these resources and become motivated to use them. Perhaps once these skills are acquired, faculty may then want to learn more about other technological tools and resources for education.

Another overwhelming response elicited from this survey was the challenges faculty encountered in this initiative regarding the time allotment and learning curves encountered so as to learn the various technologies. While this result is not surprising to anyone who has tried to utilize technology in any kind of meaningful way, it still implies that faculty are not accorded enough time to research and develop the teaching innovations which they are so strongly encouraged to use. Historically, educators have been either asked or required to incorporate a variety of teaching techniques into their pedagogy. Years of research also show that the single, most often given reason why they are not able to do this is lack of time. Rather than the results of this investigation driving the same implications that prior studies have derived (i.e. educators need release time in order to be successful with professional development), the researchers here ask not how to do this, but rather when is the education profession going to listen to the conclusions of its own research and do something about this problem?

Faculty in this research also offered their own solutions for some of the challenges that they encountered. Three participants state that faculty training and institutional support (in terms of resources) would be helpful to them. Other individuals offered suggestions that tended to imply that simply practice in thinking about and choosing content to be used with technology would help them. As this is true of any new skill, it is not surprising, although quite indicative of the skills of reflective practitioners that it is mentioned here.

When asked what technology-infused activities the faculty would like to attempt in the future, the responses given were quite ambitious. One person stated that s/he would like to teach online, while two others stated that they would like to further enhance their web development skills. Another person mentioned they would like to learn how to put their PowerPoint presentations onto a website and another individual stated that they would like to link the student portfolios to the overall assessment website. These results are quite positive because they imply that not only are faculty interested in utilizing technology, but they are also willing to challenge themselves professionally.

Subsequently, when asked what types of support these participants would need in order to be successful implementing these new activities, three people stated that support from technologists would be key, while two others desired more professional development activities. Throughout the various questions and responses on this instrument thus far, participants have made it clear that tech support and continuing education are quite important to the success of their trying to integrate technology into pedagogy. In the future, whenever technological professional development for faculty is undertaken, it seems worthwhile to plan and budget for ongoing, perhaps even individual support, for the participants. By their own indications, the rate of success in these types of projects would be much higher.

Finally, faculty members indicated that institutional support is needed in order to help them keep using technology-infused teaching and learning activities. It was emphasized that a commitment from the college to offer regular professional development workshops as well as supplying the necessary resources was key. More interestingly though, the faculty would like to see the university institute policies that "reflect its espoused beliefs" in the importance of having educators utilize technology in their pedagogy. A response such as this one indicates that although faculty may be interested in and motivated to utilize technology, despite ever-present time constraints, they do not believe that the educational institution that they work for implements policies that encourage life-long learning and professional development in their instructors.

Survey for Teacher Education Candidates

Of the 108 teacher education candidates who responded to the survey, 101 of them stated that they thought it was important for preservice teachers (such as themselves) to know how to integrate technology into teaching practices. This overwhelming response validates the mission of this initiative and others like it. With 94% of the total student sample stating that they believed it was important to know about technology integration, it seems a virtual requirement that teacher education faculty (at least at the participating institutions) be given opportunities and incentives to learn how to do this. The teacher education candidates stated that they felt this was important primarily because "education was becoming technology-dependent" and "technology is a significant component of businesses and classrooms." Ninety-five of these candidates stated that they would like to learn more about technology integration.

Ninety-six of the candidates responded that they enjoyed completing the lesson in general while 97 stated that they enjoyed using technology to complete the lesson. Ninety-two even stated that they preferred utilizing lessons that incorporated technology. While the research is yet inconclusive as to whether or not technology increases learning gains, it clearly shows that technology does not impede student learning. Given these results, and the fact that 90% of our respondents stated that they enjoyed using technology (hence it motivated them about the activity) it would make sense that university faculty would utilize technology and media in their lessons. They clearly already use non-technologic devices that may or may not increase learning, to motivate their students (i.e. humor) so why not add more pedagogical tools to the collection?

Conclusions

Based on the results of this study, it is clear that many students believe in the value of and enjoy engaging in lessons that utilize technology. As with many other pedagogical techniques or enhancements, there are also faculty who believe in and would like to learn more about the integration of technology into teaching. Convincing more faculty members to utilize technology to enhance their lessons seems less of a problem than allowing the faculty who are interested the appropriate training, development time, and resources support. As we engage in faculty development initiatives, we often tend to forget that although the goal is to get many faculty members involved, sometimes if we work with those that are truly invested in growth and improvement, in the end we reach more students. In the current project, eight faculty members initially participated but due to attrition, only four were able to implement a technology-infused lesson with their students. While these results seem less than stellar, these four faculty members were able to reach 108 university students. Perhaps the most important conclusion to be drawn from this work is that while professional development activities may not reach many faculty, the concepts do, in the end, reach many students.

References

Baldwin,G. (1999). The Evolution of Electronic Pedagogy in an outcome Based Learning Environment: Learning, Teaching and the Culture of Technology in California's Newest University – CSU Monterey Bay. ERIC Clearinghouse #IR020460

Bell, C. (1996). Managers as Mentors: Building Partnerships for Learning. ERIC Clearinghouse #ED397509

Brown, A. (1996). Bridges and Barriers to Faculty Vitality: The Grossmont College Project, 1995-1996. ERIC Clearinghouse #ED 396797

Burkett, R., Macy, M., White, J. & Feyden, C., Eds., (2001). Preservice Teacher Education. ERIC Clearinghouse #IR 020903

Clark, W.B. (1999). Professional Development for Teaching Technology across the Curriculum: Best Practices for Alberta School Jurisdictions. ERIC Clearinghouse #ED 429573

Cohen, N. (1995). Mentoring Adult Learners: A Guide to Educators and Trainers. ERIC Clearinghouse #ED 387656

Gaffney, N., Ed. (1995). A Conversation about Mentoring: Trends and Models. ERIC Clearinghouse #ED 397762

Gaston, J. & Jackson, J. (1998). Mentoring and its Implications. ERIC Clearinghouse #ED 426990

Gillan, B & McFerrin, K., Eds. (2001) Faculty Development. ERIC Clearinghouse #IR 020897

Guskey, T. & Huberman, M., Eds. (1995). Professional Development in Education: New Paradigms and Practices. ERIC Clearinghouse #394215

Guskey, T. (1994). Professional Development in Education: In Search of the Optimal Mix. ERIC Clearinghouse #ED 369181

Harnish, D. (1986). Focus on Faculty: A Report on Professional Development at Niagara County Community College. ERIC Clearinghouse #ED 286570

Henderson, J. & Welch, O. (1993). Mentoring in Higher Education and Industry: Is there a Paradox? ERIC Clearinghouse #ED 363179

Kelly, D. (1990). Reviving the "Deadwood": How to Create an Institutional Climate to Encourage the Professional Growth and Revitalization of Mid-Career Faculty in Community College. ERIC Clearinghouse #ED318518

Luna, G. & Cullen, D. (1995). Empowering the Faculty: Mentoring Redirected and Renewed. ERIC Clearinghouse #ED 399888

McAlister, K. & Reagan, C., Eds. (2001). Research. ERIC Clearinghouse #IR020906

McMullen, D., Goldbaum, H., Wolfe, R., & Sattler, J., Eds. (1998). *Using Asynchronous Learning Technology to Make the Connections among Faculty*. ERIC Clearinghouse #ED

Mihkelson, A. (1997). *A Model of Research Mentoring for Higher Education – An Overview.* ERIC Clearinghouse #ED 418661

Mullen, C. & Lick, D., Eds. (1999). *New Directions in Mentoring: Creating a Culture of Synergy*. ERIC Clearinghouse #ED 449132

Petrie, H.G., Ed., (1995). *Professionalism, Partnership, and Power: Building Professional Development Schools*. ERIC Clearinghouse #ED388628

Pierson, M., Thompson, M., Adams, A., Beyer, E., Cheriyan, S., Starke, L., Eds. (2001*). PT3 Papers*. ERIC Clearinghouse #IR020904.

Pitton, D. (2000). *Mentoring Novice Teachers*. ERIC Clearinghouse #ED 457137

Portes, P., Longwell-Grice, R. & Chan, P. (2000). *Overview of Mentoring*: Exploring Possibilities. ERIC Clearinghouse #ED 447763

Roberts & Associates. (1999). Professional Development and Learning Technologies: Needs, Issues, Trends and Activities. ERIC Clearinghouse #ED 431900

Schrum, L. & Skeele, R. (2000). Faculty Survey, Seton Hall University, PT3 Program.

Schwier, R. Akerman, M., Antifaiff, G., Arthur, P., Balbar, S. Brown, T., Byers, J., Glenn, C., Couros, A., Hansen, C., Pinder, D., Pratte, D. Scheirer, B. & Schindelka, B. (2000). The Day the Technology Died: A Tale of Murder and Redemption in a Virtual Mentorship Project. ERIC Clearinghouse #IR020445

Shaughnessy, M. (1995). Creativity in Mentoring. ERIC Clearinghouse #ED 387718

Smith, K. & Kolosick, J.T. (1996). The Shift to a Learner-Centered University: New Roles for Faculty, Students, and Technology. ERIC Clearinghouse # ED 405829

Sprague, D., Kopfman, K., & Dorsey, S. (1998). Faculty development in the integration of technology in teacher education courses. Journal of Computing in Teacher Education, 2(14), 24-28.

Strudler, N. & Wetzel, K. (1999). Lessons from Exemplary Colleges of Education: Factors Affecting Technology Integration in Preservice Programs. ETR & D, 47(4), 63-81.

Tauer, S. (1996). The Mentor-Protégé Relationship and its Effect on the Experienced Teacher. ERIC Clearinghouse #ED 397004

Valadez, J. & Duran, R. (1991). Mentoring in Higher Education. ERIC Clearinghouse #ED 331451

Ware, R. (1992). The Collaborative Role of Higher Education in the Mentor Teacher-Internship Program in New York State. ERIC Clearinghouse #ED 343866

The Evolution of Design Language Terms of Novice Instructional Designers

Kay Kyeongju Seo
Utah State University

Abstract

A design language is used to create the designer's moves, to develop each move's implications, and to construct a design process intended for a given product. This research sought to examine these functions of a design language by exploring how novice instructional designers' design language terms evolved while they were structuring an instructional product as a team. The developmental pattern of the contexts of their design language terms revealed one of the errors that novice designers could easily make: moving hurriedly to solution generation before fully analyzing the problem. This study can serve as a basis for more investigation intended to provide novice designers with guidance for executing effective design processes.

Introduction

This study was designed to explore how design language terms of novice instructional designers evolved while they were structuring an instructional product as a team. For this purpose, this study identified design language terms in the team's interactive discourse, analyzed each word in terms of its context, and drew important patterns across the data.

This study was inspired by Donald Schön's work (1983), The Reflective Practitioner. In Chapter 3, Schön defined designing as "a reflective conversation with the situation" and described how a "language of designing" between a master architect and his student shaped design events. Words fill a different role in a design process. A designer's languages are used as a vehicle to create the designer's moves, to develop each move's implications, and, ultimately, to construct a design process intended for a given product. By exploring the contexts of the team's design language terms, this study sought to examine these functions of design languages.

The underlying assumption of this study was that design languages were flexible. This characteristic of design languages was well described by Gibbons (unpublished manuscript):

> A design language is . . . a fluid collection of terms that is controlled by the designer. Over time, the language a designer uses may vary. In a sense, the evolution of a designer's style over time can be characterized in terms of changes in the designer's preferred language. (p. 12)

The findings of this study will provide the insight into the changeability of design languages that was suggested in the above statement.

Methods

Type of Research

This study was descriptive research through content analysis. The purpose of this study was to examine the progress of design language terms used by novice instructional designers and to portray the evolution of the contexts of their design language terms. To serve this purpose, descriptive research design seemed to be a logical choice. The researcher collected data from a non-experimental situation and analyzed and assembled it into a structure which would build a description of the novice designers' design language terms.

Subjects

The team consisted of five Master's-level students, two females and three males, in a simulation class of instructional technology at a western university. One of the members had some degree of knowledge of simulation design, but the rest had only minimal or no experience in creating a simulation. All male members had extensive graphic and web-design skills and were proficient with such tools as Flash, DreamWeaver, AuthorWare, FireWorks, PhotoShop, ToolBook, and ColdFusion. On the contrary, the two females had limited computer knowledge. All of the team members did their undergraduate work in fields other than instructional technology.

The Task

Students in the simulation class were instructed to complete a simulation design as a group to fulfill a course requirement. The purposes of this requirement were to provide the students with “a hands-on experience in the creation of an instructional simulation; to allow [them] to experience the simulation design process; and to give [them] a group experience in exercising practice design judgment while designing a simulation” (Course syllabus). The subject team was assigned to complete a project of creating an interactive map of Washington, DC. They were not actually required to develop an instructional simulation; they were instructed only to design a simulation and to present a prototype if possible.

Role of the Researcher

The researcher assumed two different roles in the process of data generation and data collection: One as project consultant and the other as observer. As a consultant, she provided the team with tour information on Washington, DC, and examples of various interactive maps. In addition, the researcher observed the course of the discussions and wrote fieldnotes of each meeting’s agenda.

Data Collection and Analysis

Over a two-month period from September to October 2002, the team met seven times in and out of class to discuss the design of the simulation. Each meeting lasted approximately 30 minutes. The conversations at each meeting were recorded and transcribed verbatim.

The transcriptions were analyzed as follows. First, design language terms were identified from the team’s dialogues. According to Rheinfrank and Evenson’s study (1996), design languages “consist of design elements and principles of composition . . . [D]esign languages are used to design objects that express what the objects are, what they do, and how they are to be used, and how they contribute to experience” (p. 68). All words or phrases in the transcriptions that fit this description were selected as design language terms.

Next, each word was examined in terms of its context and categorized according to a scheme that was adapted from Rowland’s categorization scheme (1991). Rowland categorized the design procedures of a team of instructional designers under 17 operations including Read, Rephrase, Describe, Record, Retrieve, Question, Identify, Infer, Predict, Generate, Evaluate, Reason, Conclude, Review, State doing, Plan, and Goal. For this study, the researcher used 15 out of 17 operations and modified definitions of some of the operations. Table 1 shows the revised scheme used in this study and examples of how a design language term was categorized under one of the contexts.

Table 1 *Context Categorization Scheme*

(Adapted from Rowland’s categorization scheme developed for encoding the protocols)

Contexts	Definitions and Examples (Underlined are design language terms)
Read	Read directly from materials Ex: “This is what I have. ‘It includes basic elements of each aspect.’”
Rephrase	Repeat in own words something recently read; say what is being recorded Ex: “You mean ‘implemented’”.
Describe	Comment the appearance or contents of an object or situation Ex: “And then they go into some type of cost.”
Retrieve	Recall knowledge or experience from memory Ex: “She showed us the model we want to do”
Question	Pose a question Ex: “Are we including landmarks?”
Identify	Label something as a problem, constraint, or resources without certainty Ex: “Almost like an interactive storyboard.”

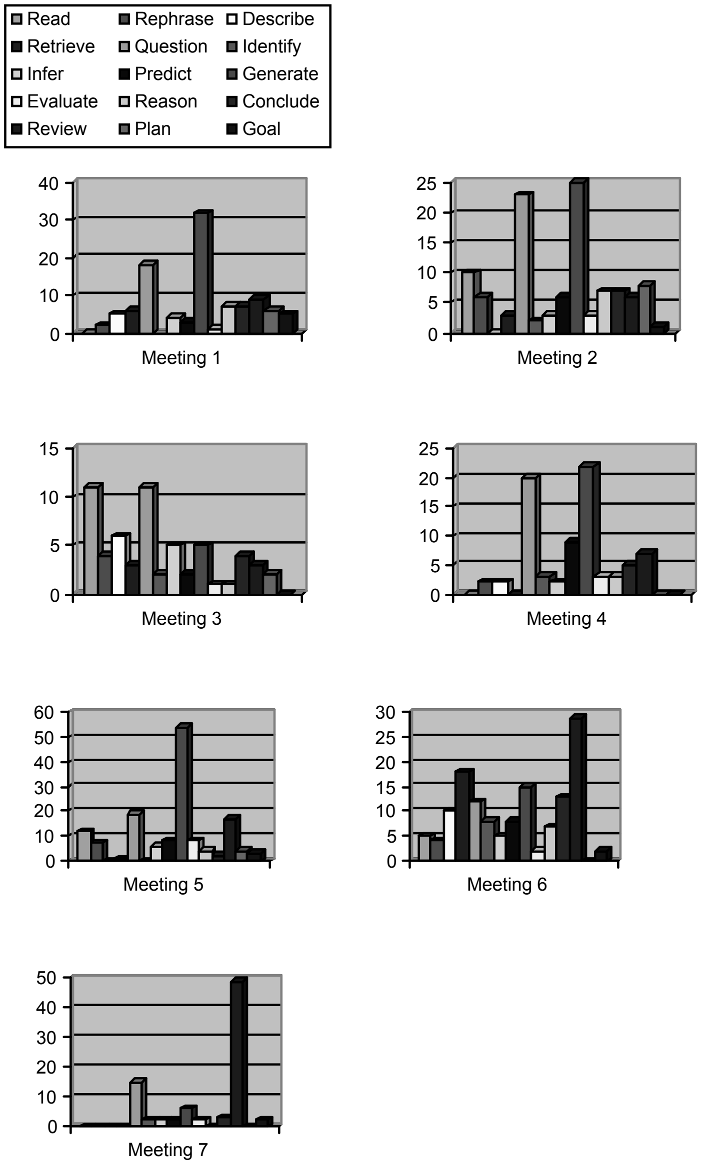
Infer	Derive an aspect or detail Ex: "There's got to be some <u>instructional objectives</u> for that somewhere."
Predict	Foretell or guess the result(s) of an action Ex: "If we <u>decipher the map</u>, we will run into the same problems."
Generate	Propose how a problem may be solved, create some aspect of a strategy Ex: "You could add a <u>real world</u> context to that."
Evaluate	Access information, understanding, solution potential, task difficulty, etc. Ex: "I don't know how it's <u>utilize</u>d."
Reason	State why something is, should be, or will be Ex: "…cause they are on the same <u>tracks</u>."
Conclude	State with certainty, verify, or confirm Ex: "They have a <u>standard map</u>."
Review	Summarize understanding or progress toward goal, solution, etc. Ex: "So essentially, we are <u>instruct</u>ing them how to use it."
Plan	State intention to do something Ex: "What we need to do now is figure out the user's <u>problems</u>."
Goal	An end to be worked toward or aimed at Ex: "The goal is to help them be able to <u>interact with the map</u>."

The number of each operation that occurred at each meeting was then counted and represented in a graph. Seven graphs from seven meetings were compared to track how the contexts evolved.

Results

The seven graphs from Figure 1 illustrate how the contexts of the design language terms evolved throughout the seven meetings. The first graph shows that the novice designers started to generate ideas as early as in the first meeting. Many questions were posed, and as many solutions were proposed in the second meeting. The graph from the third meeting shows that the bar labeled "read" became notably taller. According to the researcher's fieldnotes, the team reexamined what they had discussed by reading their design document in this meeting. The following meeting went back to the interplay between questions and solution generation. From Meeting 5, the number of times the item "question" was used decreased remarkably compared to the "generate" bar. We can assume that many of the problems started to be resolved after the proposed solutions of this meeting. And finally, the team concluded the design process by reexamining the design in the seventh meeting: As you can see, the bar "review" was tallest in this meeting.

Figure 1. Transition of the Contexts of the Design Language Terms



Two major characteristics of the transition of the contexts were identified from the analysis of the dynamic change in the seven graphs. First, a context category that was most highly used across the data was the item "generate." This category's bar was tallest in Meetings 1, 2, 4, and 5. Also, the second most frequently used category was the context "question." The item "question" followed "generate" as the second tallest bar in Meetings 1, 2, 4, and 5, and it outnumbered "generate" in Meetings 3 and 7. From the definitions of these two contexts, it can be accurately concluded that the novice designers primarily focused on problem solving by "posing a question" and "proposing how a problem may be solved" throughout the whole series of meetings.

Second, implied in the above characteristic is that the team did not pay much attention to understanding what the problem was and determining what goals they wanted to achieve, which should have been considered in the early phases of design. Figure 1 shows that the occurrence of the items such as "evaluate" and "plan" was very low in all seven meetings. This lack of problem analysis caused a serious predicament after the last meeting, by which time the team was convinced that they had successfully completed the design of an interactive map. It turned out that their design was different from what the instructor had proposed in the beginning. Unfortunately, they had to revise a considerable part of the design.

Conclusion

This study is important to the field of instructional technology. The developmental pattern of the contexts of the design language terms reveals one of the errors that novice instructional designers easily make: Moving hurriedly to solution generation before fully analyzing the problem. Without accurate, in-depth problem assessment, one cannot produce a product that effectively resolves the problem. By describing the thought processes of the novice instructional designers, this study serves as a basis for more investigation intended to provide novice designers with guidance for executing effective design processes.

Topics for further research should include design languages of expert designers. It would be interesting to see how the nature and contexts of experts' design languages are different from novices'. Narrowing the gap between experts' design languages and novices' design languages will significantly contribute to the successful implementation of effective and efficient design processes.

References

Gibbons, A. S. (Unpublished manuscript). Chapter 4: Design languages of instruction. Retrieved September 5, 2002, from http://eres.usu.edu/tempfiles/tmp21589/Chapter_4__2002_Design_Languages_of_Instruction.pdf

Rheinfrank, J. & Evenson, S. (1996). Design languages. In T. Winograd (ed.), *Bringing design to software*. Reading, MA: Addison-Wesley Publishing Company.

Rowland, G. (1991). What do instruction designers actually do?: An initial investigation of expert practice. *Performance Improvement Quarterly, 5*(2): 65-86.

Schön, D. A. (1983). The reflective practitioner: How professionals think in action. New York: Basic Books

Use of CoWebs In Scenario-Based ID Instruction

Neal Shambaugh
West Virginia University

Abstract

This paper documents the use of CoWebs in a graduate-level instructional design (ID) course to facilitate online peer review of design decisions incorporated in students' ID projects. CoWebs are shared web pages, which allow for and depend on user-created structure and content. Student feedback examines not only design decisions, but the use of scenarios, envisioning the educational intervention. Students reflect on these scenario descriptions and subsequently generate iterative responses to an instructional problem. are presented. The paper reports an analysis of CoWeb postings, ID phase submissions, and student perceptions of the CoWeb activity from a Fall, 2003 implementation. Implications for the use of CoWebs in ID courses are provided.

Background

When asked to identify the instructional problem underlying an educational intervention, student designers tend to propose, even "fixate," on a solution without a clear understanding of the instructional problem (Shambaugh & Magliaro, 2001). The only people who seem disturbed by this behavior are probably instructional design instructors. Within ID models, critique of design decisions is represented by "dotted lines" or boxes denoting "revision" or "formative evaluation." Periodic or benchmark reviews in real development settings may be limited to specific technical features of the work without standing back and appraising the overall potential of the design to address learning. Our conscious need to make decisions frequently takes precedence over thinking about action. Such a move may be due to a human propensity to proceed to a solution in light of existing information (Simon, 1996) and the desire to act to decrease uncertainty, even survival (e.g., negotiating city traffic). In the case of instructional problems is this human trait appropriate?

Norman (1993) cautiously suggests two types of cognition to consider here: experiential, the dynamic, real-time thinking necessary to keep us from getting hit by a bus; and reflective, a mode of thought that suggests possibilities, considers options, and makes decisions. Rowland, Fixl, and Yung (1992) have already suggested the value of reflectivity in ID instruction, but how do we balance the need to think about what we do, while not paralyzing us from making timely decisions, even a less than optimal decision given limited information?

A recent pedagogical study (Shambaugh, 2002) documented the use of scenarios to take advantage of this human propensity to move quickly to resolution and to address the frequent de-coupling of reflection (e.g., formal design reviews) from design activity (Schön, 1983). Designer-constructed scenarios, tapping existing experience and information, provided condensed descriptions of proposed solutions to instructional problems (Carroll, 2000). Scenario descriptions are nothing more than conversations or written paragraphs that record a tentative approach to a problem or need. This study reported that scenarios helped students to identify and clarify the instructional problem, identify learning needs prior to needs assessment, and contributed to a more focused and motivated needs assessment activity.

As Norman (1993) suggests, "modern technology has the power to enhance reflection and to make it more powerful than before" (p. 16). To augment the in-class scenario activity and peer conversations away from class, online CoWebs were chosen. CoWebs are collaborative web pages, which are easily edited and may include text, links, visuals, and other media elements. The developers of CoWebs used them in an architectural studio and found that the environment "failed to promote open interaction and did not appear to sustain a strong sense of community" (Craig & Zimring, 2000, p. 1). Reasons for this failure were cited as the individual and competitive nature of architectural studios, the inability of CoWebs to provide the "finished" look and 3-dimensional features of architectural design prototypes, and insufficient navigational mapping of pages. However, CoWebs were shown to support the ill-defined activity of early design ideas. The authors also cited that the use of text on the CoWeb pages was not particularly useful for architectural designers, as these students favored complex visuals to represent design decisions.

Use of Collaborative Web Pages

Collaborative web pages, CoWebs, are sometimes referred to as Wiki's. A Wiki, according to the

TLT Group, is “half of a Hawaiian term for quick” and denotes a collaborative website comprised of the collective work of many authors (http://www.tltgroup.org). Wiki’s have been in use since 1995 and are popular with scientists, software engineers, and others collaborating on projects. Characteristics of Wiki sites include the following (http://c2.com/cgi-bin?WhyWikeWorks):

- Anything posted can be changed by anyone.
- “Anyone can play, but only good players last” as “low signal” content can be easily removed.
- “An intelligence test of sorts,” as use is relatively easy but requires a little time to figure out how edit, upload and interlink pages.
- Does not work in real time and thus is suitable for reflectivity. Pages require time for people to think, write, and edit.

In summary, features that may be viewed as flaws, are attributes for participation and reflectivity, inviting people who want to be there and have a space where people want to have a real dialogue. Thus, the appeal factor is low. Ward Cunningham, who developed the first Wiki in 1995, lists the major attributes of Wiki’s and collaborative web pages (http://www.tlgtroup.org/CommunityConnectedness/AsynchTools.htm):

- Open: any user can edit any entry. This includes adding content, turning existing content into links, or changing content altogether. Involves trust, a sense of collectivity.
- Organic: the structure and content of the site evolves as long as it’s used. Users continually define extent of site.
- Mundane: easily learned coding conventions, such as building links, formatting text, uploading images.
- Incremental: entries can cite other entries, including pages that have not been written yet. Encourages interconnection.
- Precise: entry titles encourage focus, highlighting of crucial terms and concepts.
- Observable: activity within the site can be watched and reviewed by any visitor to the site.

The current technical status of CoWebs, while limiting to other designers, seems to lend themselves to ID instruction. The mode of representation in ID projects has been primarily textual, using tables and narrative with some limited use of visuals to represent design decisions. The environment of ID instruction is considerably less competitive than any architectural design studio. The overall focus for this study was “In what ways can CoWebs help student instructional designers to reflect on and revise design decisions?” Specific research questions included the following:

4 Question 1: Did the CoWeb pages improve clarity of the instructional problem?
5 Question 2: Did the CoWeb pages help students study their instructional problem using needs assessment?
6 Question 3: Did peer feedback on the CoWeb pages contribute to revised instructional design decisions?
7 Question 4: What were student perceptions of the CoWeb activity?

Methodology

Course Design Sequence

In the master’s level instructional design course, which was the context for this study, students chose an instructional problem and designed a response to this educational need. Shambaugh and Magliaro (1997) was used as the text, which presents to students an ID process consisting of Context, Analysis, Design, Prototype, and Evaluation phases. The principal assessment tool is an ID project, representing students’ ID decisions. A set of 15 Design Activities structures and guides students’ ID thinking decision-making, and reflection. A web board provided “conferences” for student posting of all work (see http://anatomy.hsc.wvu.edu:8080/~edp640). The web board also stored course materials (syllabus, design activities, resources, research papers). The second web presence included the Collaborative Web Pages, which was located on a student’s server. As these pages can be easily accessed for anyone who knows the web address, a URL is not provided here.

The first two phases in the ID process, Learning Beliefs and Design Tools (principally ID models), were used to establish the context for the traditional ID process. Design-A-Lesson and Learning Principles tasks helped students to reflect on how they currently plan instruction and their views on learning. Students drafted a Mission Statement, which was used to assess how students' learning beliefs were being applied in their projects. Students also sketched a visual of their own ID model and provided an explanatory narrative.

The subsequent Needs Assessment phase structured students' research into the instructional problem and options for addressing it. A personal conference provided individual assistance with their needs assessment strategy in terms of what to study, who to talk with, what references to consult, and how to summarize their findings. Based on Needs Assessment, students identified goals for their project and revised their original design intent. Students were then introduced to design phases, which included Instructional Sequence, Assessment, Instructional Framework, Instructional Media/Technology, and a Prototype component of the instructional intervention (e.g., unit, web pages, tutorial, workshop activities). During the Instructional Frameworks phase in which teaching options were specified, students demonstrated a teaching strategy in their Prototype. Instructional Media/Technology was addressed throughout the course beginning with Needs Assessment. Students were prompted to be open to media and technology possibilities and to make a case for how their choices supported their goals. A second personal conference addressed Program Evaluation and project issues. During the final week of the course students revised their personal ID models and submitted a written self-evaluation of the course and their learning.

Collaborative Web Pages

Students identified an instructional problem and recorded their initial understanding of the problem through a written Intent Statement. Prior to conducting a needs assessment, students were asked to write a scenario envisioning a successful implementation of the project. The scenario description and instructional problem statement were posted on individual CoWeb pages. Students were organized into pairs or triads, based on their project type. Students were then prompted to provide online feedback to their peer designers.

The home page for the CoWeb site included two links. The Peer Review page listed links to students CoWeb pages organized by pairs or triads, depending on content area. A second link pointed to Individual Designer's pages, a location provided for optional post personal and professional images, text, and links. Students could easily add text to their individual pages or respond to peer pages by selecting an "EDIT" icon at the top of the CoWeb interface screen, typing in their comments and selecting "SAVE." A "HELP" option provided students with details on changing font size and color, insertion of dialogue boxes (eliminating the need to use Edit/Save), insertion of media files, and URL's.

Data Sources And Analysis

Participants in this study included 15 students enrolled in a master's level instructional design course. Eleven students identified Instructional Technology as their major or field of interest. Three other students identified their majors as Educational Psychology, Curriculum and Instruction, and Special Education. One student was an instructor in Foreign Languages. Collaboration, using the CoWeb pages, was organized around similar ID projects. The groups were labeled as Cognition, Literacy, Pedagogy, Software Use, Tech Training, Technology Literacy, and Languages.

Data collected for this study was limited to the first 7 weeks of the Fall, 2003 semester, in order to meet the deadline for this paper submission. Three data sources were used to answer the four research questions: (a) student postings on CoWeb pages, (b) student design submissions, and (c) student perceptions of the CoWeb task. CoWeb postings included an Initial Project statement, which included a statement of the instructional problem, a Revised Scenario Description (initial version was conducted in class), Needs Assessment questions, a Revised Intent Statement, and Peer Feedback.

The analysis of the CoWeb postings and ID decisions used what is called participatory analysis. The data consists of students generating features and consequences in a scenario; students questioning stages of action at each phase ID; and systematic questioning of each event, act, or intent in a description (Carroll, 2000) by the instructor. Participatory analysis examines the CoWeb postings and ID decisions to the degree to which peers provided feedback that showed up in students' project revisions.

A table for each student was used to display transcribed text from the CoWeb postings and Design Submissions (Miles & Huberman, 1988). The table consisted of three columns, including CoWeb Posting, Design Submission, and Peer Feedback. Each student's table consisted of four rows, recording text for Problem Statement, Scenario Description, Needs Assessment Questions, and Revised Intent Statement. Using this table, data reduction employed a second table to record judgments on how each student (a) improved clarity of the instructional problem from initial Problem Statement to Revised Intent Statement (Yes/No-Comments), (b) whether the scenario description was examined in the needs assessment (Yes/No-Comments), and (c) whether peer feedback contributed to design revisions (Yes/No-Comments). A frequency count of Yes and No tabulations was reported. This data reduction table served to record data analysis answering research questions 1-3.

Collecting student perceptions of the CoWeb task (research question 4) was accomplished in Week 7 by asking students their reactions to the CoWeb task ("How might these pages better help you in your project?). These comments were transcribed, coded for "Yes, they helped," "No, they did not help," or "Unsure," as the question was asked during Week 7 of the semester. A frequency count of Yes, No, and Unsure tabulations was reported. Any suggestions for CoWeb improvement were noted and summarized.

Results

Summary of the data reduction and frequency counts for the four research questions are reported in Figure 1. A total of 15 students were involved.

Research Question	Yes	No	Unsure
Did the CoWeb pages improve clarity of the instructional problem?	9	6	
Did the CoWeb pages help students study their instructional problem using needs assessment?	12	3	
Did peer feedback on the CoWeb pages contribute to revised instructional design decisions?	9	6	
What were student perceptions of the CoWeb activity?	2	4	4

Figure 1. Frequency Responses of Data Reduction

Instructional Problem

Research question 1 asked if the CoWeb pages improved clarity of the instructional problem? The initial instructional problem statement was recorded on the CoWeb page with feedback from peers, and after a needs assessment, was changed in the Revised Intent Statement. Of the 15 students, 9 students were determined to have improved clarity of the instructional problem statement. Clarity was interpreted partly as a clearer sense of instructional need. Examples of educational needs addressed by these problem statements included "increasing parental involvement in children's literacy development," "lack of Lotus Notes skills," "lack of appreciation for music theory." Clarity improvement was also interpreted as a clearer sense of instructional problem choice. For example, one student initially viewed "the lack of a college course" as the problem, but reframed the problem by making a case for why students would need a proposed course. Another student viewed senior citizens as lacking in computer skills, but in a revised Intent Statement, reframed the problem as "helping senior citizens make connections to family and information." A third example was incorporating the student's "love for literature" to reframe a view of "old" literature (the view from students) as "classic" literature (the view of teachers who want to instill this love in students). A third interpretation of instructional problem clarity was the length of the problem statement. Based on ten years of involvement with this course, by the end of the course a student's problem statement is encouraged to be one sentence long. However, by Week 7, some of the problem statements were still long. Five of the 15 problem statements were shorter and one was actually longer.

Needs Assessment

The second research question asked if the CoWeb pages helped students to examine their instructional problem in a needs assessment? Twelve of the 15 students were helped by the CoWeb questions and peer responses. The analysis for this question examined the scenario description posted on the CoWeb page, the peer responses to the scenario, and what students wrote in their needs assessment findings. To see the improvements across different instructional problems, Figure 2 lists the accomplishments of the needs assessment activity. One of the more significant accomplishments was one student who originally proposed the increased use of creativity in public school lessons. The outcome of her needs assessment was to reframe her problem to helping teachers “teach to all students” and to raise teacher awareness to learner differences.

YES/NO	Needs Assessment Findings
YES	Developing a home environment to help children write, speak, and listen
YES	Investigated cultural appreciation activities; pilot test of activity in local library
YES	Interviewed wide range of people
NO	Scenario not done
YES	Interviewed students on multimedia course needs and software choices
YES	Clear listing of content to be learned; studying adult learner differences
NO	Scenario not done
YES	Investigating teaching strategies to improve appreciation of learning music theory
YES	Listing open-source software learning activities
YES	Researching media options for second language learning
YES	Understanding the relevance of computers to senior citizens
YES	Project change from creativity in lessons to helping teachers teach to all students
NO	Problem choice narrow and defined – filling out forms and using software
YES	Discovered integration of second language learning in courses an innovative idea
YES	Realized need for teacher buy-in of literature teaching strategies and use of workshops

Figure 2. CoWeb Use Contribute to Needs Assessment Findings

Revised ID Decisions

The third research question asked if peer feedback on the CoWeb pages contributed to design revisions. Of the 15 students, 9 were helped by the CoWeb page postings and feedback. The data sources here included the CoWeb page postings, feedback, and needs assessment submissions. Although 9 out of 15 students were evaluated as being assisted by the CoWeb postings, this was based on text evidence that comments on the pages were showing up as design decisions. However, only a low level of feedback and suggestions was provided. Suggestions included “conduct a survey,” “OK to your teaching strategy,” and prompting for detail or clarity on some details.

Student Perceptions

The fourth research question asked for students’ perceptions on the CoWeb task. After seven weeks into the semester only two students believed the CoWeb pages were helping them, four students said “No,” the pages were not helping them, and 4 students wrote that they were “unsure” as to their value. Five suggestions were offered for their improvement and they included: allow all students to provide feedback, rather than just the groupings; allow groups to form along student interests; provide more structure to the CoWeb site, structure the CoWeb site to parallel the project requirements, and provide more user-friendliness to the pages.

Implications for ID Instruction

Limitations of Research

One limitation to this pedagogical study was including only seven weeks of data. This was necessitated by the semester schedule and the Proceedings deadline. However, sufficient data and student comments provided me with evidence as to how the CoWebs were helping or were not helping. Another limitation is one that characterizes study of all instructional interventions, and that is the contextual influence of other instruction. Particularly, in this study, feedback to students came from multiple sources:

students in class, students' postings on the CoWeb pages, teacher-to-student via email, in-class, and feedback on student submissions. A third limitation was the necessity to interpret performance based on what students posted, what students said, and what students submitted. Ten years of experience with this course has helped to ensure some measure of trustworthiness in interpretation of "clarity of instructional problem," for example. Suggestions for improvement in this research require improvements in the structure of the CoWeb task within the context of learning instructional design.

Improvements For CoWeb Page Use

This research confirms what many instructors already know; namely, that students tend to value the feedback of the instructor over other students. Email was cited by one student as more preferred. Although no reason was cited, the more personal and private form of exchange may have been the reason. One of the attributes of collaborative web pages is their openness to viewing, commentary, and revision. Students expressed in class, as well as in their brief comments, some uncertainty with this innovation. When faced with a simple interface that does not look like a typical software graphical-user interface, students are unsure of what to do. Although "Help" and other icons provide an interface for students to learn how to contribute to CoWeb pages, they tend to "stop" when faced with a blank screen. Initially, I provided the structure that guided students to groups and to individual pages, as well as providing type-in dialogue boxes to facilitate initial writing.

My next iteration of using CoWeb pages are to use them to post their weekly submissions, keyed to the ID project, and it is here that I will provide online feedback. In this way, the instructor is directly involved in commenting on student work, not through physical paper, but via the CoWeb pages. A class environment will continue to be used, but commenting on student work will use the CoWeb structure. One comment here, however, is that this structure, which students look for in instruction, somewhat defeats the purpose of the open and organic potentials of CoWeb sites.

To facilitate peer review and peer design, the peers must be invested in the design of an actual project. This authentic use of the CoWebs has already been demonstrated by pair programming used by computer programmers (Baer, 2003, September). This could be accomplished during a summer delivery of the ID course, where groups would develop an ID project. Another option would be for online delivery where the online presence becomes crucial for feedback. For the near term, instructor-student use of the CoWebs may be a more efficient strategy than the one used in this study. Peer review could be structured along the ID project submissions and students could be required to post critiques during the design stages, such as a critique of student's instructional sequences (story boards, lists, maps). Teaching and assessment, and the use of instructional media and technology could also be critiqued using a formal task structure.

Improvements For ID Instruction

The combination of the online Web Board and the CoWeb pages helped to speed-up needs assessment by about one week. Students posting their weekly work to the Web Board allowed me to comment on work by the day it was due. The CoWeb pages gave students additional feedback, ideas, and prompting. In the past ten years of instruction, usually four weeks were provided for Needs Assessment activity. This allowed more design time during the semester.

This paper continues my ongoing study at improving ID instruction and ultimately ID use. A previous Proceedings paper (Shambaugh, 2002) examined the use of the scenario description in ID instruction and how these descriptions help students to envision, reconsider, and revise their design within the context of actually designing; thus, speeding up the design-reflect cycle and ultimately, better and more timely instructional designs (Carroll, 2000). The next phase of this development continues the use of scenario descriptions, CoWebs as an online presence for design activity, and reconfiguring design activity as scenario building, a process I hope becomes useful for actual ID use.

References

Baer, M. (2003, September). The new X-men. *Wired*, 124-129.

Carroll, J. M. (2000). *Making use: Scenario-based design of human-computer interactions*. Cambridge, MA: MIT Press.

Craig, D. L., & Zimring, C. (2000). Supporting collaborative design groups as design communities. *Design Studies, 20*(2), 187-204.

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Readings, MA: Addison-Wesley.

Rowland, G., Fixl, A., & Yung, J. (1992, December). Educating the reflective designer. *Educational Technology*, 36-44.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Shambaugh, N. & Magliaro, S. G. (1997). *Mastering the possibilities: A process approach to instructional design*. Boston, MA: Allyn and Bacon.

Shambaugh, N. & Magliaro, S. G. (2001). A reflexive model for teaching instructional design. *Educational Technology Research and Development, 49*(2), 69-92.

Shambaugh, N. (2002). A Scenario Strategy to Promote Design Task Reflection. *2002 Proceedings of the American Educational Communications and Technology Annual Conference*. Dallas, TX: AECT.

Simon, H. A. (1996). *The sciences of the artificial (3rd ed.)*. Cambridge, MA: MIT Press.

Understanding and Facilitating Historical Argumentation Skills in a Multimedia Learning Environment Among High School Students

Namsoo Shin
University of Michigan

Steven McGee
John A. Baro
Wheeling Jesuit University

Abstract

The Center for Educational Technologies® is developing Foundations of Freedom™, a DVD-based multimedia learning environment (http://www.cet.edu/constitution). Foundations of Freedom™ helps high school students explore the development of the U.S. Constitution and essential constitutional issues throughout history. This study sought to understand students' argumentation skills in the domain of constitutional history. It then identified developmental areas in their argumentation abilities. The results show that to argue successfully, students need to improve at discussing multiple perspectives, obtaining historical evidence, and identifying the limitations of their own decision. Various learning tools in Foundations of Freedom facilitate those weak components.

Argumentation is essential in social science, especially history, for promoting students' critical thinking and decision making. National standards for historical thinking stress the importance of students creating a historical argument of their own (National Center for History in the Schools, 1996). Historical argumentation can be defined in the following way, "a position is supported by offering historical evidence appropriate in a particular situation" (Cerbin, 1988). The essential processes of a historical argument include (a) analyzing and interpreting of issues and events; (b) considering the issues and events from multiple perspectives; (c) obtaining historical evidence to make warranted decision; d) evaluating the documents for the implementation of a decision; and e) alternative courses of action (National Center for History in the Schools, 1996; Toulmin, Rieke, & Janik, 1984; Voss, Blais, Means, Greene, & Ahwesh, 1986). Students need to coordinate representations of several documents, resolve any inconsistencies among documents, and weigh a information in light of its sources for developing a successful argument in history (Rouet, Britt, Mason, & Perfetti, 1996). The need for developing instructional strategies that improve argumentation skills in history is widely recognized. Many educators have developed instructional activities and teaching methods for improving students' historical argumentation skills (Doppen, 2000; Drake, 2002; Risinger, 1992; Swan, 1994; Tunnell & Ammon, 1996; Wineburg, 1991).

Yet, the results of research indicate that the current education has little impact on students' argumentation skills in history (Cavalli-Sforza, 1991). The reports of the 2001 National Assessment of Education Progress in U.S. history show that a high percentage of students in K-12 did not reach the advanced level achievement, which requires historical argumentation skills (Patric, 2002). Perkins (1985) found that students' ability to construct arguments is generally disappointing and does not significantly improve from high schools to graduate schools. Additionally, high school students have little knowledge of creating historical interpretation, although they may know a lot of history (Doppen, 2000).

Students' shortcomings in the argumentation skills may be attributed to the scarce opportunities for argumentative activities in high school education. Students do not have much experience at facing multiple perspectives. In history classes, students encounter an oversimplified account of history instead of considering various sources of evidence (Tunnell, & Ammon, 1996). In addition, textbooks usually present a single perspective with few conflicting ideas; and a narrow view of history with little variation in the interpretation of our past (Foster, & Rosch, 1997). Secondly, students in schools do not provide various source documents nor do they learn the importance of the documents in the course of developing their arguments (Wineburg, 1991). Lastly, teachers are forced to cover too much material with too little time and consequently focused on content rather than process (Doppen, 2000; O'Neill, 2003). The intention of this study is to create a rich multimedia learning environment in order to overcome the deficiencies in the current history classes. The investigation was conducted as a task analysis to specify the weak components for developing a successful argument among high school students in history. Based on the results of the study, various scaffolding tools in Foundations of Freedom were designed to facilitate those weak components.

Method

Participants

The participants in this study were 28 ninth-grade history students. For conducting think-aloud protocol, three additional subjects were selected: a ninth-grader, a junior college student, and an expert in history. There were 17 females and 14 males. All were Caucasian.

Instrument

The Center for Educational Technologies team constructed an open-ended essay to understand students' argumentation skills on historical and constitutional issues (Cerbin, 1988). The open-ended question was, ***"Defenders of the War Powers Resolution praise it as consistent with our system of government based on separated powers. In contrast, the resolution's opponents maintain that it undermines that principle. What is your position on this issue?"*** In this task students addressed one theme, separation of power in the U.S. Constitution. They had to compose in a historical context their own claim to the constitutional issue of the War Powers Resolution.

The team provided students with historical reading, including the U.S. Constitution, the Gulf of Tonkin Resolution, the Senate debate record on Tonkin, an article on the invasion of Cambodia, President Nixon's speech on Cambodia, and the War Powers Resolution.

Scoring Systems

Student papers were scored according to the criteria from a Toulminian argumentation theory and from historical argumentation (Toulmin, Rieke, & Janik, 1984). A Toulminian argumentation theory defines the important elements in an argument: a claim, grounds, a warrant, a backing, a qualifier, a rebuttal (Toulmin, Rieke, & Janik, 1984). The rubric dimensions contain five features: a claim, multiple perspectives, evidence, the limitations of one's own decision, and an opposing position. Here's what the criteria focused on for each of the five features:

- Claim—how well does the argument state a clear position relating to the essay's theme.
- Multiple perspectives—how many perspectives relevant to the topic does the argument state and how well are the perspectives integrated to support the claim.
- Evidence—how relevant is the historical evidence to support the claim and how well is the evidence integrated within the argument.
- Limitations of own decision—how well are the limitations or drawbacks of the student's decision identified and how well does the student defend the decision in the face of the limitations.
- Opposing position—how effective is the student in evaluating and comparing the weaknesses and strengths of the counter position with his or her own decision.

The criteria provided specific descriptive information for students' argumentation processes. The scores ranged from 0 (low) to 4 (high) for a claim and evidence. The scores ranged from 0 (low) to 3 (high) for multiple perspectives, the limitations of the decision, and an opposing position. Two raters evaluated the compositions independently. The average of the two scores was used in the analysis. Overall interrater reliabilities, determined using Cronbach Alpha, was .92.

Procedures

Students first were given reading a day before as a homework assignment. They were allowed to spend as much time as necessary to read the documents. The students then received the open-ended essay in class. Additionally, investigators tape-recorded the responses of the three aforementioned "think-aloud" subjects. They were asked to think aloud while they tackled the essay. After finishing, they were interviewed individually to verify how they approached the question and what they thought the task was about. All participants were allowed to use the reading documents as references when answering the question. They also were given as much time as necessary to finish. Most were done within 45 minutes.

Data Analysis and Results

The data came from written and oral argumentation tasks. Student responses were scored according to the rubric systems and analyzed using descriptive statistics and think-aloud protocol methods. Argumentation skills were investigated by examining whether students formulated a claim, discussed multiple perspectives, provided evidence, stated limitations of their own decision, and evaluated opposing positions in composing an historical argument. Student responses were rated for the quality of their claim, the number of perspectives, the evidence, the statement of limitation, and the discussion of opposing positions using descriptive statistics and think-aloud protocol methods. The total scores were correlated to the five components to understand a pattern of students' arguments. Table 1 shows means, standard deviations, intercorrelations, and interrater reliabilities for each of the five components of students' arguments.

Table 1. *Means, Standard Deviation, Interrater Reliabilities, and Intercorrelations of the Variables with the Total Scores*

Trait	Mean	SD	Correlations	Interrater Reliabilities
Formulates Claim	2.45 (61%)	1.00	.64**	.86
Discusses Multiple Perspectives	1.5 (50%)	.54	.25	.91
Obtains Evidence	.48 (12%)	.82	.15	.91
Identifies Limitations of Own Decision	.50 (17%)	1.04	.50*	.98
Evaluates Opposing Position	1.8 (60%)	1.03	.70**	.91

Note. ( ). Standardized proportion of the means. **. $p < .01$. *. $p < .05$.

According to the results, the students scored 61 percent ($\underline{M}$ = 2.5 out of 4) in formulating the claim, 50 percent ($\underline{M}$ = 1.5 out of 3) in discussing multiple perspectives, and 60 percent ($\underline{M}$ = 1.8 out of 3) in evaluating the opposing position. Students scored less than 20 percent in obtaining evidence ($\underline{M}$ = 0 .48 out of 4) and identifying the limitations of their decision ($\underline{M}$ = 0.5 out of 3).

The intercorrelations (Pearson Product Moment Correlation) of three features—formulating a claim ($\underline{r}$ = .64), identifying the limitations of their decision ($\underline{r}$ = .50), and evaluating the opposing position ($\underline{r}$ = .70)—were statistically significantly correlated with the total scores. However, there were not statistically significant relations in discussing the multiple perspectives and obtaining the evidence.

Additionally, the differences of argumentation processes were analyzed by comparing those of an expert, a college student, and a high school student. In the analysis of think-aloud protocol, the responses of the subjects were categorized into the five components by examining (a) which components the subjects presented and (b) how the subjects represented the components in their arguments. Then each component was compared individually among the three subjects to understand the differences in the argumentation patterns among an expert, an intermediate, and a novice in history. Although the analysis is preliminary, several noteworthy findings arose in the analysis of think-aloud protocol. The expert concentrated on finding relevant evidence and considering all possible perspectives before reaching a conclusion. In contrast, the ninth-grader decided a position immediately after reading the background information and the question. This person tended to focus on brief descriptions in the background information rather than on available historical evidence to develop an argument.

Discussion and Implications

The study showed that the argumentation skills of students in the target high school can be considered poor in history. The results indicated that students have moderate skills in formulating a claim and evaluating the opposing position for developing a historical argument. The results further demonstrated that the weaknesses are in discussing multiple perspectives, obtaining historical evidence, and identifying the limitations of one's own decision.

In the criteria of discussing multiple perspectives, the mean of 1.5 shows that students focused on only one or two out of three possible perspectives related to the topic. Moreover, the results of correlation highlighted that the constructs of discussing multiple perspectives did not discriminate between good and poor students. It is speculated that most of students are able to identify directly presented perspectives but not deeper-level perspectives in a context. On the basis of the results, students are needed to encourage for searching all possible perspectives in a context to develop a successful argument.

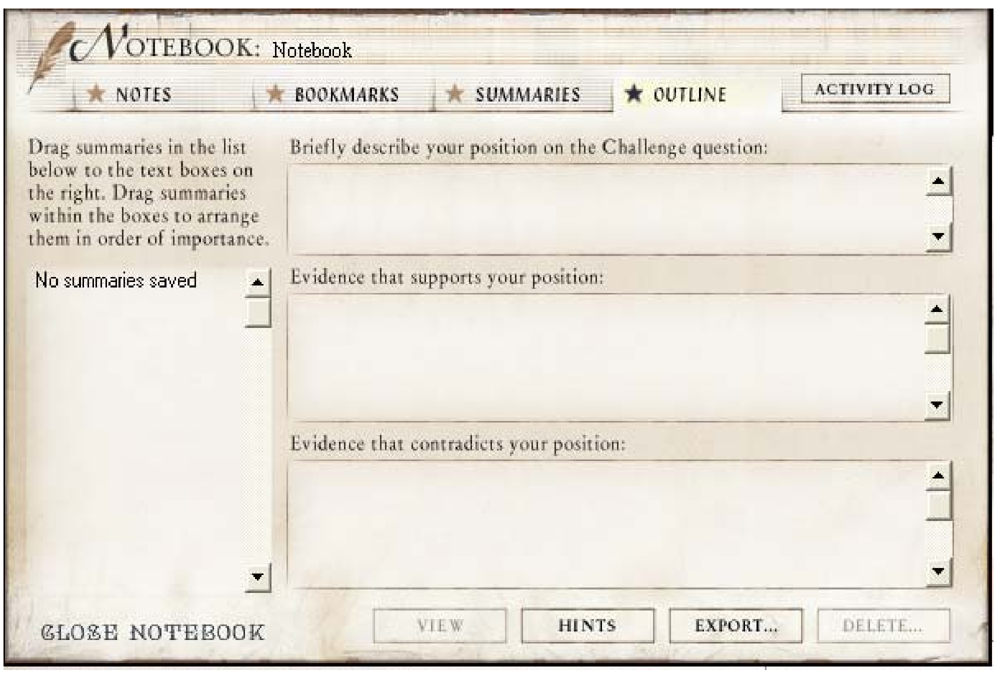
In the case of discussing multiple perspectives, the correlation of the obtaining evidence with the total scores did not discriminate between good and poor students. Additionally, the mean proportions of the scores were less than 15 %. The results indicate that obtaining evidence from sources is the most difficult component in the development of an argument for high school students. The results are consistent with previous research studies (Crammond, 1998; Wineburg, 1991). Designers must put considerable effort toward scaffolding students to collect evidence from a variety of sources.

The students who developed a successful argument were proficient in identifying the limitation of their own decision. However, most students failed to identify the limitations of their own decision based on the students' low mean scores. It is necessary to support students for identifying the limitation of own decision in order to support the development of a successful argument.

Finally, the results of think-aloud protocol demonstrated that the expert spent an enormous amount of time on planning, finding possible perspectives, and selecting relevant evidence. Then they formulated a claim and evaluated the opposing position. In contrast, the high school student focused mainly on formulating claims and evaluating the opposing position.

The results suggest that instructional designers should create a learning tool to scaffold students' weak components: discussing multiple perspectives, obtaining relevant historical evidence, and stating the limitations of their decision. The development team for Foundations of Freedom created learning tools to scaffold students' weak components. The Archive—a search engine—was designed to encourage students to collect evidence from multiple perspectives (See http://www.cet.edu/constitution/archive.html). The Bookmarks and Summaries were designed to facilitate students to obtain relevant evidence from available documents in the Archive (See http://www.cet.edu/constitution/notebook.html). Finally, the Outline tool was designed for scaffolding students to state the limitations of their decision (See Figure 1). The main principle was encouraging students to collect evidence for all possible claims (positions) from the Bookmarks and Summaries. The representation Outline tool provides students opportunities to consider the limitation of their decision because they are able to see the evidence for opposite decisions.

Figure 1. Foundation of Freedom: Outline



References

Cavalli-Sforze, V. (1991). *An Environment for Tutoring Argumentation Skills.* In working notes of the AAAI'91 spring symposium series: Argumentation and belief, 71-84. Stanford, CA.

Cerbin, B. (1988). *The Nature and Development of Informal Reasoning Skills in College Students.* Paper presented at the National Institute on Issues in Teaching and Learning, Chicago, IL.

Crammond, J. (1998). The uses and complexity of argument structures in expert and student persuasive writing. *Written Communication, 15*(2), 230-268.

Drake, F.D. (2002). *Teaching Historical Thinking.* Bloomington, IN: Clearinghouse for social studies/social science education. ED 467383.

Foster, S.J., & Roaxh, R. (1997). Teaching World War I from multiple perspectives. *Social Education, 61*(7), 429-434

National Center for History in the Schools. (1996). *National Standards for History.* Los Angeles: National Center for History in the Schools.

O'Neil, D.K. (April, 2003). *Sharing Accountability through Sharing Our Accounts: Piloting an on-line community for high school history learning.* The Annual Conference of American Education Research Association. Chicago, IL.

Patric, J.J. (2002). *The 2001 NAEP in U.S. History.* Bloomington, IN: Clearinghouse for social studies/social science education. ED 465707.

Perkins, D.N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77*(5), 562-571

Risinger, C.F. (1992). *Trends in K-12 Social Studies.* Bloomington, IN: Clearinghouse for social studies/social science education. ED 351278.

Rouet, J., Britt, A.M., Mason, R.A., & Perfetti, C.A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology, 88*(3), 478-493.

Swan K. (June, 1994). *History, Hypermedia, and Criss-Crossed Conceptual Landscapes: Designing Hypermedia Applications to Support the Development of Historical Thinking.* Educational Multimedia and Hypermedia. Vancouver, British Columbia, Canada.

Toulmin, S., Rieke, R., & Janik, A. (1984). *An Introduction to Reasoning.* New York: Macmillan.

Tunnell, M.O., & Ammon, R. (1996). The story of ourselves. Fostering multiple historical perspectives. *Social Education, 60*(4), 212-215.

Voss, J. F., Blais, J., Means, M. L., Greene, T. R., & Ahwesh, E. (1986). Informal reasoning and subject matter knowledge in the solving of economics problems by naive and novice individuals. *Cognition and Instruction, 3*(4), 269-302.

Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83,* 73-87.

Effects of Goals-Supported Learning Environments and Peer Group Composition on Motivation and Problem Solving

Hae-Deok Song
Barbara L. Grabowski
Penn State University

Abstract

Learners' goal orientation is an important factor to influence motivational efforts in a problem-solving environment. Despite the importance of goal orientation as a motivation variable, the implications of goal orientation for the instructional design practice have not been thoroughly investigated. The purpose of this study is to examine the effects of goal-structured learning environments with peer group composition on motivation and problem solving. The results of this study indicated that learning goal-structured environment had a significant effect on motivation while heterogeneous peer group composition had a significant effect on monitoring and evaluation skill, one of problem solving skills. This study suggests that learners' motivation can be enhanced in a problem solving environment by the learning goal orientation support.

Introduction

Researchers have recognized that one of the main elements required for the successful problem solving is motivation (Albanese, 1993; Berkel & Schmidt, 2000; MacKinnon, 1999; Schmidt & Moust, 2000). Even though the function of motivation in problem solving has gained considerable attention, it is not clear which specific motivation variables influence problem solving ability (Mayer, 1998).

One such motivation variable, goal orientation, is currently receiving much attention in the field of motivation research. Goal orientation means desire or purpose that individuals possess when they pursue competence in achievement situations (Harackiewicz, 1998; Pintrich, 2000). Therefore, without considering learners' goal orientation, it is difficult to motivate students. According to goal orientation theory, there are two primary contrasting goal types: learning and performance goals. Research shows that a generally adaptive pattern of outcome is associated with learning goals and a generally less adaptive pattern is associated with performance goals (Meece, 1991; Pintrich, 2000). However, it is not clear what type of goal orientation is positively related to motivation and problem solving.

Early researchers tended to view goal orientation as static personal disposition. Recently, however, some researchers have suggested that goal orientation can be affected by contextual features and have investigated the effects of individual contextual information designed to increase goal orientation, such as environmental cues (Elliot & Dweck, 1988; Gabrielle & Montecinos, 2001; McNeil & Alibali, 2000; McWhaw & Abrami, 2001; Schunk, 1996). However, a common limitation of these past studies is that they have failed to holistically take into account a combination of the elements that affect goal orientation. To fully support goal orientation, it may be necessary to design a learning environment that includes more holistically essential elements influencing goal orientation. In addition, given the mixed effects of the two types of goal orientation on student outcomes, it is important to determine which types of learning environment support may lead to high motivation and problem solving.

It is believed that goal-orientation supports mediate the effectiveness of independent variables. Once goal-structured learning environments are provided in a problem-solving environment, they have an effect on students' perception of peer group environment goal structure and goal orientation. It is also expected hat these will mediate the effectiveness of motivation and problem solving skills. However, studies of goal orientation have showed mixed results for the relationship between goal-structured learning environment and goal orientation, between goal orientation and motivation, and between motivation and problem solving skills.

Besides the importance of supporting goal orientation, it is also clear that the peer group learning can facilitate motivation and problem solving (Ames, 1992; Epstein, 1989; Meece, 1991; Maehr & Midgley, 1991; Pintrich & Schunk, 2002). To have effective peer group work, group composition should be addressed. Typically, peer group learning advocates recommend that students be grouped heterogeneously

(Hooper et al., 1993). One such factor related to motivation is self-efficacy because when students see themselves as efficient learners, they are highly motivated. Thus, it is necessary to examine the effects of heterogeneous group according to self-efficacy level in a problem solving environment.

Therefore, the purpose of this study was to examine 1) the effect of two types of goal-structured learning environments and 2) the effect of types of peer group composition according to self-efficacy levels on the perception of peer group environment goal structures, goal orientations, motivation, and problem solving skills. Interaction effects between these two variables were also investigated.

The following specific research questions are formed:

1. Do the learning goal-structured environment and the use of heterogeneous peer group have an effect on motivation?
2. Do the learning goal-structured environment and the use of heterogeneous peer group have an effect on the problem solving skill?
3. Is there a correlation between perception of peer group environment goal structure, goal orientation, motivation, and problem solving skill?

Theoretical Framework

The Role of Goal-Orientation in Problem Solving

One of the main components required for successful problem solving is motivation. Without a strong motivation to commit themselves to learning, students will not be able to acquire higher problem-solving skills (Schmidt & Moust, 2000). Although the importance of motivation has long been recognized, it is not clear which motivation variables affects problem solving. A limited number of studies on the role of motivation in problem solving have been conducted, to date. Thus, it is necessary to find effective motivation variables and to define the motivation mechanisms of those variables. Goal orientation is one motivation variable that can explain gains in academic achievement. They explain why students engage in achievement behavior. A goal, in particular, is regarded as an integrated pattern of belief that leads to different ways of engaging, acting, and responding to achievement situations.

Goal Orientations and Links to Motivation and Problem Solving

Researchers have agreed that there are two types of goal orientation and they have labeled these two goals *learning (mastery, task-involved) goal orientation* and *performance (ego-involved, ability-focused) goal orientation* (Ames & Archer, 1988; Dweck & Leggett, 1988; Elliot & Dweck, 1988; Maehr & Midgley, 1991; Nicholls, 1984). A learning goal orientation involves a focus on learning, mastering tasks, trying to gain understanding, and improving competence according to self-imposed standards. A performance goal orientation, on the other hand, has a focus on demonstrating ability, trying to surpass normative performance standards, striving to be the best in the group, seeking public recognition for high-level performance, and avoiding judgment for low ability.

According to the studies on the effects of goal orientation and motivation carried out by Rawsthorne & Elliot (1999), a learning goal orientation is positively related to task enjoyment, a classical measure of motivation. Most researchers who believe in the effects of learning goals on motivation have claimed that performance goals undermine intrinsic motivation by leading learners to focus on demonstrating superior ability rather than on getting involved in the task itself. Recently, however, some researchers have reported that performance goals do not always lead to less interest, motivation, or task involvement. For example, Harackiewicz et al. (1998) suggest that the effects of both learning and performance goals depend on students' personal characteristics and on the context of the activity and thus, both types of goals can increase motivation. Therefore, it is necessary to investigate the relationships between different goals and motivation, to discover what type of goals may be most effective.

Although goal-orientation theorists have extensively reviewed the relationships between goals and cognitive outcomes, limited research has been conducted, to date, on how goals are linked to the use of problem-solving strategies (Ames & Archer, 1988; Dweck & Leggert, 1988). According to Pintrich & Schunk (2002), this is an area on which future research should focus in order to define the nature of goal orientation. However, considering that the main components of problem solving are the ability to represent a problem, the ability to develop a solution, and the ability to evaluate or monitor that solution, one can assume that goal orientation is also related to problem-solving skills, because those skills are related to the use of the cognitive strategies discussed above: deeper processing of cognitive strategies, self-monitoring strategies, and self-regulatory strategies.

Enhancing Goal Orientation: Goal-Structured Learning Environments

Researchers have suggested that the goal orientation an individual adopts may be influenced by contextual information or situational factors. They have claimed that goals are not stable personal characteristics but rather are very sensitive to contextual features. Accordingly, they have suggested that goals can be changed through contextual features. Grounded on the view of goals as a state, several studies have examined the effects of contextual information on outcomes (Elliot & Dweck, 1998; Gabriele & Montecinos, 2001). However, one limitation of these studies is that they don't systemically take into consideration the elements that affect goal orientation. Since a goal is an integrated cognitive representation, one must take various factors into account in order to fully develop or change goal orientation.

What elements should we consider in order to support goal orientation? To design a learning environment that helps students to develop goal orientation, it is necessary to identify essential elements that influence two types of goal orientations; learning goal orientation and performance goal orientation. Previous studies reports that three essential elements also foster both learning and performance goal orientation (Blumenfeld, 1992; Maehr & Midgley, 1993; Meece, 1991; Roeser et *al*., 1996): (a) task, (b) authority, and (c) recognition or evaluation. However, little research has compared the effects of learning environments differently structured toward different types of goal orientations. While the literature on goal-orientation theory generally implies that a learning goal-structured environment has a positive effect, and that a performance goal-structured environment has a negative effect, the results have, in fact, been mixed. For example, Fuchs et *al*. (1997) empirically tested the effects of a classroom environment that was structured to facilitate learning goal orientation, and found that three elements (task, authority, and evaluation) affect students' achievement. However, other investigations have also reported positive effects of the performance goal-structured environment (Butler, 1992; Harackiewicz et al., 1998; Sansone, 1986). Therefore, it is still not clear which type of goal-structured learning environment may be effective.

Enhancing Goal Orientation: Group Composition

Research on peer group learning has shown it to be effective in increasing students' goal orientation (Nichols & Miller, 1994). To facilitate the development of such an orientation, researchers suggest that groups should not be rigidly set but rather fluid and adaptive (Martin, 2001). Heterogeneous grouping is one example of an effective group composition that will foster a learning goal orientation. According to Johnson & Johnson (1984), the heterogeneous grouping provides affective and cognitive benefits to both high-and low-ability learners. Within heterogeneous collaborative groups, less able learners can benefit from emulating more able learners' attitude, while more able learners can benefit from explaining their knowledge structures to less able learners. In addition, the motivation of less able learners is likely to be increased, since they will receive more attention in a heterogeneous group than in a homogeneous group (Hooper & Hannafin, 1988).

Some researchers on cooperative learning have suggested that factors other than ability level should be considered in forming effective groups. Given the interest in motivation in this study, a desired personal characteristic factor should be related to motivation perspectives. One possible personal characteristic to be considered is self-efficacy. *Self-efficacy* is particularly important as a motivation construct because it mediates the relationship between goals and performance. For example, research on goal orientation has demonstrated that students with learning goals also rate themselves high on self-efficacy (Pintrich, 2000). Given the correlation between goal orientation and self-efficacy, it is also believed that self-efficacy may predict the adoption of learning goals (Linnenbrink & Pintrich, 2002). Considering the link between self-efficacy and goals, self-efficacy would appear to be an appropriate personal characteristic to be taken into account in forming effective groups. However, little research on group composition by self-efficacy has been conducted, as yet. Therefore, it is necessary to vary group composition according to self-efficacy.

Method

Participants

Subjects were 90 sixth-grade students from four classrooms at two rural middle schools located in the northeastern United States. A total of 96 students participated in the beginning of the study. However, 90 participants in the final analyses were analyzed due to missing data.

Research Design

A 2 x 2 randomized factorial design was used. The primary independent variables in this study were types of goal-structured learning environments (learning goal, performance goal) and peer group composition (heterogeneous, homogeneous). Students were assigned to one of the following four treatment groups resulting from a crossing of two factors: (a) learning goal-structured/ heterogeneous peer group, (b) learning goal-structured/ homogeneous peer group, (c) performance goal-structured/ heterogeneous peer group, and (d) performance goal-structured/ homogeneous peer group. Dependent measures were motivation and problem solving skills.

Experimental Treatments

Goal-Structured Learning Environments. Two different versions of web-based PBL treatment materials structured differently toward either learning goals and performance goals were used in this study. Web-based PBL tutorials have been adapted and developed from one of the lesson plans in a supplementary web-enhanced PBL science curriculum, Kids as Airborne Mission Scientists (KaAMS). The main problem of the lesson is "What is your flight Plan for Airborne Mission and Why?.

Each treatment included three essential areas prompting achievement goals: a) Task, b) Authority, and c) Evaluation (Ames, 1992; Fuchs et al, 1997). First, students in Learning Goal-Structured Environment (*LGSE*) (a) received task message stressing the importance of challenging work and intrinsic value of learning, (b) provided choice control such as setting priorities, and (c) provided self-referenced evaluation information. The aim of providing the *LGSE* was to promote a learning goal-oriented climate in a peer group. Second, students in Performance Goal-Structured Environment (*PGSE*) (a) received task message stressing the importance of performance, (b) provided instructor-designed choice, and (c) provided norm-referenced evaluation standard and social comparison information. The aim of providing the performance *PGSE* is to promote a performance goal-oriented climate in a peer group.

The treatment materials were reviewed for content and face validity by a faculty member in the program of Educational Psychology and they were pilot tested for readability and interpretation with eight students from one classroom who did not participated in this study during the pilot test period.

Peer Group Composition. High self-efficacy students scored at or above the median, which fell at the 51st percentile, and low self-efficacy students scored below the median in the pretest. In the heterogeneous group, one high and one low self-efficacy level subject were assigned. In the homogeneous group, two high or two low self-efficacy level subjects were assigned.

Assessment Instruments

Five instruments were used in this study.

Math and Science Self-efficacy: Fouad et al.'s (1997) assessment of math/science self efficacy (12 items) was used because it was developed for measuring middle school students' self-efficacy.

Perception of Peer Group Goal Environment Structures: Perceived Peer Group Goal Structures was used from the Manual for the Patterns of Adaptive Learning Scale (PALS) (Midgley *et al.,* 2000). There are 17 peer group goal structure items (Learning-6, Performance-11). Both learning and performance goal structures of peer group environment were assessed with a 5 point Likert- type scale ranging from 1 (Not at all true) to 5(Very true).

Goal Orientation: Personal Goal Orientation was used from the PALS (Midgley et al., 2000). There are 14 personal goal orientation items (Learning-5, Performance-9). Each learning and performance goal orientation is assessed with a 5 point Likert- type scale ranging from 1 (Not at all true) to 5(Very true).

Motivation: A science subscale was selected and adapted from Children's Academic Intrinsic Motivation Inventory (CAIMI) (Gottfried, 1985). 24 items out of 26 items in the original scale were assessed in this study with a 5 point Likert Scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Problem solving skills: The problem-solving reports written by individuals in each group were analyzed in terms of three components of problem solving using a scoring rubric. The scoring rubric was adapted from

Ge's (2001) rubric system for the solving of an ill-structured problem. Performance was measured according to degree of quality on the three major constructs of problem solving: problem representation, solution development, and monitoring and evaluation of problem solving.

Procedure

A self-efficacy pretest was administered to all participants in the target classrooms to identify students with either high or low levels of self-efficacy two weeks before the experimentation. Based on group assignment by self efficacy, the study was administered in three separate sessions with 45 minutes for each session: (a) reviewing the problem scenario and defining the problem, (b) investigating the problem and proposing solutions at a group level, (c) presenting the solution at an individual level. Students completed a different version of the treatment structured toward a different goal orientation. Students in *LGSE* solved a problem with the treatment designed to facilitate a learning goal orientation; students in *PGSE* solved the problem with the treatment designed to foster a performance goal orientation. The next day after the treatment period, students completed a package of survey questionnaire that measured perception of peer group goal environment structure, goal orientation, and motivation.

Data Analysis

Evaluation of Motivation: To investigate the effects of goal-structured learning environment and peer group composition on motivation, a factorial analysis of covariance (ANCOVA) was applied. The factorial ANCOVA was selected in this study because self-efficacy is a covariate.

Evaluation of Problem-Solution Skills: To identify the effects on problem solving skills, two data analyses were applied: a factorial Multivariate Analysis of Covariance (MANOVA) was conducted to test for the effects on problem representation and solution development and univariate Analysis of Variance (ANOVA) for the effect on monitoring and evaluation.

Relationship between Peer Group Goal Structure, Goal Orientation, Motivation, and Problem Solving Skills: Pearson's r was employed to find the correlation between mediating variables (peer group goal structure, goal orientation) and dependent variables (motivation, three components of problem solving skills).

Results

Effects of Goal-Structured Learning Environment and Group Composition on Motivation

First, the factorial ANCOVA showed a main effect for the types of goal-structured learning environments. Students working in the learning goal-structured environment (M=3.49, SD =.69) had a significantly higher motivation score than the students working in the performance goal-structured environment (M=3.16, SD =.66), $F\ (1,\ 89)$ = 3.97, p <.04, η^2 =.05.

Second, the factorial ANCOVA showed no main effect for the types of peer group composition. Students working in the heterogeneous peer group (M=3.38, SD =.65) had not a significantly higher motivation score than the students working in the homogeneous peer group (M=3.28, SD =.75), $F\ (1,\ 89)$ = .20, p >.05, η^2 =.00.

Third, the results of the factorial ANCOVA did not reveal the interaction effect between types of goal-structured learning environments and types of group composition, $F\ (1,\ 89)$ = .04, p >.05, η^2 =.00, and thus, failed to support the hypothesis. Despite the result, students working in the learning goal-structured environment and the heterogeneous peer group achieved the highest motivation score (M = 3.49, SD = .69) among four treatment groups.

Effects of Goal-Structured Learning Environment and Group Composition on Problem Solving Skills

First, in order to determine the effects of goal-structured learning environment and peer group composition on *problem representation* and *solution development,* a factorial Multivariate Analysis of Covariance (MANCOVA) was conducted. The results of factorial MANOVA did not reveal significant effects for both main effects (types of goal-structured learning environment, type of peer group composition) as well as interaction effect between two treatments.

Second, in order to determine the effects of goal-structured learning environment and peer group composition on *monitoring and evaluation*, an univariate analysis of ANOVA was conducted. The results of univariate ANOVA did not reveal a significant difference for the types of goal structured learning environment, $F(1, 86) = .09$, $p = .76$, $\eta^2 = .00$. However, there was a significant main effect on monitoring and evaluation skill between types of peer group composition, $F(1, 86) = 5.57$, $p < .05$, $\eta^2 = .06$. Students working in the heterogeneous group had an overall average of .57 ($SD = .96$) on monitoring and evaluation skill while students working in the homogeneous group had an overall average of .18 ($SD = .50$). No significant interaction was found between types of goal-structured environment and types of group composition, $F(1, 86) = .58$, $p < .05$, $\eta^2 = .01$.

Correlations between Perception of Peer Group Environment Goal Structure, Goal Orientation, Motivation, and Problem Solving

Correlations in Learning Goal-Orientation Path Variables

Pearson's *r* was calculated between perceived learning goal structure, learning goal orientation, motivation, and problem solving skills for students who participated in learning goal-structured environment and heterogeneous group, respectively.

First, Pearson's *r* was calculated for students (N=46) who participated in learning goal-structured environment. Significant correlations were found (1) between the perceived learning goal structure and the learning goal orientation (r=.789, $p <$ .01) and (2) between the learning goal orientation and motivation (r=.388, $p <$ .01). These results showed that students who perceived the learning goal structure in their learning goal-structured environment tended to have higher learning goal orientation and motivation scores. These suggest that an increased learning goal orientation as a result of participation in the learning goal-structured environment might enhance motivation.

Second, Pearson's *r* was calculated for the students (N=46) who worked in heterogeneous peer group, Significant correlations were found (1) between the perceived learning goal structure and the learning goal orientation (r=.639, $p <$ .01), (2) between the perceived learning goal structure and motivation (r=.327, $p <$ .05), (3) the perceived learning goal structure and solution development (r=.354, $p <$ .05), (4) between the learning goal orientation and solution development (r=.394, $p <$ .05), and (5) between motivation and solution development (r=.353, $p <$ .05, N=46). These results revealed that students who perceived the learning goal structure in their heterogeneous peer group tended to have higher scores in the learning goal orientation test, motivation test, and solution development skill test. Students with the learning goal orientation also tended to have higher motivation scores. These suggest that perceived learning goal structure in the heterogeneous peer group might influence motivation as well as one of problem solving skills, solution development.

Correlations in Performance Goal-Orientation Path Variables

Aside from the correlations in the learning goal orientation path variables, Pearson's *r* was calculated between perceived performance goal structure, performance goal orientation, motivation, and problem solving skills for students who participated in performance goal-structured environment and homogeneous group, respectively.

First, Pearson's *r* was calculated for the students (N=44) who worked in the performance goal-structured environment. Significant correlations were found (1) between the perceived performance goal structure and the performance goal orientation (r =.603, $p <$ .01), (2) between performance goal orientation and problem representation (r = -.314, $p <$ .05), and (3) between motivation and solution development (r = .338, $p <$ .05). These results indicated that the students who perceived the performance goal structure in their performance goal-structured environment tended to have higher performance goal orientation scores. Interestingly, a significant negative correlation was found between the performance goal orientation and the problem representation (r = -.314, $p <$ .05, N = 44). These results indicated that students who had the performance goal orientation tended to have lower scores in problem representation.

Second, Pearson's *r* was calculated for the students (N=44) who worked in the homogeneous group. Significant correlation was found between perceived performance goal structure and the performance goal orientation (r =.573, $p <$ .01). This result indicated that the students who perceived the performance goal structure in their homogeneous group environment tended to have higher performance goal orientation scores. Interestingly, there was also a negative correlation between the performance goal orientation and motivation (r = -.140, $p >$ .05), between the performance goal orientation and the problem representation (r = -.219, $p >$ .05, N = 44), and between the performance goal orientation and solution development(r = -.206, $p >$ .05, N = 44). These results indicated that students who had the performance goal orientation tended to have lower scores in the motivation test as well as at two components of problem solving skill tests: problem representation and solution development.

Conclusion

The finding on significant effect of learning goal-structured environment on motivation supports the hypothesis that learning goal orientation support better leads to higher motivation scores than performance goal orientation support. It corroborates empirically the research by Rawsthone & Elliot (1999) that learning goal orientation is positively related to motivation.

The results of peer group composition on monitoring and evaluation skill, one of problem solving skills support the hypothesis that group composition provides cognitive benefits to both high and low levels of self-efficacy students. This finding extends previous studies results by Johnson & Johnson (1984) that heterogeneous grouping allows group members to examine the problem from several different perspectives at a problem solving environment.

The correlation results showed significant correlations between perceived goal structure and goal orientation. These results confirm the research approach contending that learning goal orientation is very sensitive to context and one's learning goal orientation can be changed through contextual factors (Epstein, 1989; Maehr & Midgley, 1991). Interestingly, a significant correlation was found between learning goal orientation and motivation whereas a negative correlation was found between performance goal orientation and problem representation. These results indicate the effect of learning goal orientation over performance goal orientation on motivation and problem representation.

The findings of this study have important implications on the design of problem solving environments that help students to increase their motivation. Given that motivation can be supported by a combination of the elements that affect learning goal orientation, it is important to structure learning environments based on these elements to motivate learners in problem solving environments. Future study is also needed to examine the effect of each element in learning goal-structured environment. This will help to identify the most effective element in structuring environment toward learning goal orientation.

References

Albanese, M., A., & Mitchell, S. (1993). Problem-based learning: a review of literature on its' outcomes and implementation issues. *Academic Medicine, 68*(1), 52-81.

Ames, C. (1984). Achievement attributions and self-instructions under competitive and individualistic structures. *Journal of Educational Psychology, 76*(3), 478-487.

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261-271.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation process. *Journal of Educational Psychology, 80*(3), 260-267.

Berkel, H., & Schmidt, H. G. (2000). Motivation to commit oneself as a determinant of achievement in problem-based learning. *Higher Education*, 231-242.

Butler, R. (1992). What young people want to know when: Effects of mastery and ability goals on interest in different kinds of social comparisons. *Journal of Personality and Social Psychology, 62*(6), 934-943.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

Elliot, E. S., & Dweck, C. S. (1988). Goals: an approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*(1), 5-12.

Epstein, J. (1989). Family structure and student motivation: A developmental perspective. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 259-295). San Diego, CA: Academic Press.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, G. L., Katzaroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal, 34*(3), 513-543.

Gabriele, A., J., & Montectinos, C. (2001). Collaborative learning with a skilled peer: The influence of achieving goals and perceptions of partners' competence on the participation and learning of low-achieving students. *The Journal of Experimental Education, 69*(2), 152-178.

Ge, X. (2001). *Scaffolding students' problem solving processes on an ll-structured task using question prompts and peer interaction.* The Penn State University.

Harackiewicz, J. M., Barron, K., & Elliot, E. S. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist, 33*(1), 1-21.

Hooper, S., & Hannafin, M. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Computing Research, 4*(4), 413-424.

Hooper, S., Temiyakarn, C., & Williams, M. D. (1993). The effects of cooperative learning and learner control on high- and average-ability students. *Educational Technology Research and Development, 4*(2), 5-18.

Johnson, D. W., & Johnson, R. T. (1984). Cooperative small group learning. *Curriculum Report, 14*(1), 2-7.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: an asymmetric bidirectional model. *Journal of Educational Psychology, 37*(2), 69-78.

MacKinnon, M. J. (1999). CORE elements of student motivation in problem-based learning. *New Directions for Teaching and Learning, 78*(Summer), 49-58.

Maehr, M. L., & Midgley, C. (1991). Enhancing student motivation: a school wide approach. *Educational Psychologist, 26*(3&4), 399-427.

Martin, V. (2001). *The effect of a mastery motivational climate motor skill intervention on student achievement and behavior in a naturalistic physical education setting.* Unpublished Ed. D., Auburn University.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science, 26*, 49-63.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: externally imposed goals influence what is learned. *Journal of Educational Psychology, 92*(4), 734-744.

McWhaw, K., & Abrami, P. C. (2001). Student goal orientation and interest: Effects on students' use of self-regulated learning strategies. *Contemporary Educational Psychology, 26*, 311-329.

Meece, J. L. (1991). The classroom context and students' motivational goals. *Advances in motivation and achievement, 7*, 261-285.

Midgley, C. (Ed.). (2002). *Goals, goal structures, and patterns of adaptive learning*. New Jersey: Lawrence Erlbaum Associates.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E. M., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the patterns of Adaptive Learning Scale*. Ann Arbor: University of Michigan.

Nicholls, J. G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91*(3), 328-346.

Nichols, J. D., & Miller, R. B. (1994). Cooperative learning and student motivation. *Contemporary Educational Psychology, 19*, 167-178.

Pintrich, P. R. (2000). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology, 25*, 92-104.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in Education-Theory, Research, and Applications* (2nd Edition ed.). New Jersey: Merrill Prentice Hall.

Rawsthrone, L. J., & Elliot, E. S. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review, 3*(4), 326-344.

Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology, 88*(3), 408-422.

Sansone, C. (1986). A question of competence: The effects of competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology, 51*(5), 918-931.

Schmidt, H. G., & Moust, J. H. (2000). Factors affecting small-group tutorial learning: A review of research. In D. H. Evensen & C. E. Hmelo (Eds.), *Problem-based learning : a research perspective on learning interactions* (pp. 19-51). Mahwah, N.J.: L. Erlbaum Associates.

Schunk, D. H. (1996). Goal and self-evaluative influences during children's skill learning. *American Educational Research Journal, 33*(2), 359-382.

Music, Art and Multi-Media: Arranging Time and Space

David H. Stuart
Iowa State University

Abstract

Two scholars at Iowa State University have combined their disciplines of music and art to develop a series of lectures to present basic elements of form. The approach uses examples of locally available works of art and a variety of music from different cultures and time periods to acquaint students with the techniques of repetition, contrast and variation and our predictable human responses of anticipation, suspense, delay of satisfaction, heightening of tension, and surprise. Using locally available works of art encourages students to seek out the originals, and using music from many cultures, traditional western-European art music as well as contemporary popular music -- enables us to respect both the musical and artistic tastes of our students. The "tag-team" approach -- an expert from each of the areas alternating music and art, technique and response -- helps students better understand the universality of these formal elements.

Introduction

Every semester, more than 620 students take Introduction to Music Listening at Iowa State University, and nearly 100 are turned away and put on a waiting list. The course has gained the reputation as one which doesn't restrict its coverage to the music of dead, white European males, but which includes non-Western music, rock, jazz, folk, and the latest popular music. Students also know that they will not only *hear* music, they will *see* it. At the department's state-of-the-art Martha-Ellen Tye Recital Hall, huge screens project video recordings of recitals, concerts, and demonstrations, while surround-sound satisfies the most sophisticated and technically savvy young music lover.

I have long understood the necessity of visualizing music. There was a time, before radio, and before the advent of a flourishing recording industry, when listeners were entertained exclusively by live musicians. The visual and the aural were not and could not be separated. Society's hunger to "return" to the luxurious and fulfilling aesthetic experience of combining the two senses in the late twentieth century was well illustrated with the explosive popularity of television's rock videos. Whenever possible, therefore, I "show" my students what they are listening to. Videos run the gamut from Maestro Sergiu Celibidache rehearsing Prokofiev's *Classical Symphony*, to Branford Marsalis and his jazz trio, to the Muungano National Choir of Kenya singing the *Missa Luba*.

This technique proved so successful, that I began to wonder if I couldn't illustrate abstract musical concepts visually. Iowa State University is home to one of the nation's largest art-on-campus collections, and a museum staff devoted to educating students about the collection and the concepts they represent. In her discussions of visual literacy, Lynette Pohlman, director of Iowa State University Museums, does much the same thing that I do in my music listening class. Additionally, many of the visual concepts she covers are identical to the aural ones I include. Both art and music manipulate the basic human emotions of *satisfaction, anticipation, suspense, delay of satisfaction, and surprise* by almost identical techniques: *repetition, contrast, and variation.*

Music and art are distinctively personal disciplines. The passion that viewers and listeners feel for their art and music may be entirely unrelated to any intellectual processing or deep understanding. Their passion relates to their personal experiences, past exposure, and peer groups, in addition to formal education and training. The more that teachers of both art and music can gain the trust of their students, the more successful they will be in introducing new concepts, new music, and new art. Trust can never be gained if teachers dismiss or devalue their students' passions or experiences. Trust is gained by taking seriously all that goes into a person's individual passions.

There is an important added bonus of using art on campus, or local art, to illustrate abstract concepts. A student may walk past a campus sculpture on his way to class every day. She may pass a painting at the library, or a stained glass window at church. By using everyday experiences to illustrate concepts, and by respecting both the musical and artistic tastes of our students, we will be closer to true teaching and learning.

Together, Lynette Pohlman and I use the distinctive art of Iowa State to illustrate concepts common to the visual arts and music, making use of the paintings and sculptures of regionalist painter Grant Wood and sculptor Christian Petersen. The student reception was so positive that when I was invited to teach at National Taiwan Normal University in spring, 2003, I used the same approach. But, of course, instead of using Iowa regionalist artists, I used artwork that my students could visit at the Taipei Palace Museum.

What follows is a description of basic concepts of music, illustrated by the art that is central to the experience of students at National Taiwan Normal University. The same approach can be used in Omaha, Madrid, or Toronto. And the concepts are true for the Beetles or the Who, Haydn or Vivaldi, and Beijing and Italian opera. For the sake of presentation, I will pair the "techniques" and the "emotional responses," but any of the responses can and often do accompany any of the techniques.

Repetition

Repetition can be used on both a small and large scale. Let's look at a work of art and two musical examples that use repetition in different ways to achieve harmony and balance. In a detail from *Leaf 4, Washing the Buddha* from *Spring Morning in the Han Palace*, the artist uses short lines, both horizontal and vertical, to create an elegant balance throughout the work. These short lines function as tiny building blocks -- short lines for the roof lines, similar lines used to demarcate the walk ways, and vertical lines that change direction to form the back of the screen and the pillars within the main building.

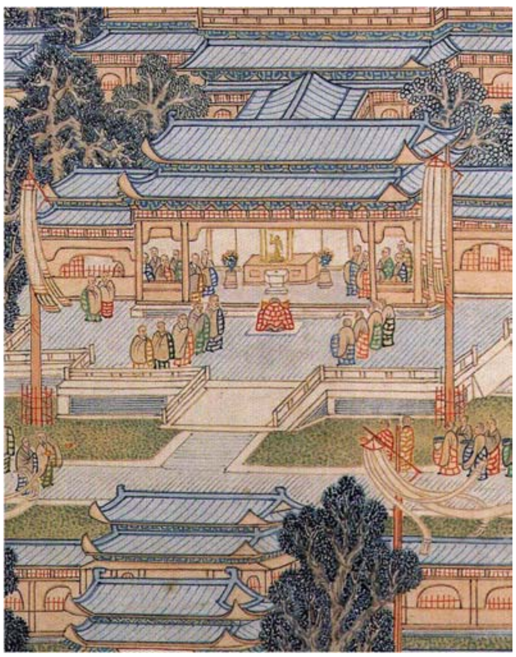


Figure 1. Detail from *Leaf 4, Washing the Buddha* from *Spring Morning in the Han Palace*

The artist also repeats horizontal and vertical lines on a larger scale to balance the whole structure. The horizontal lines of the walkways and roof tiles expand to become the rooflines in both horizontal directions to demarcate the elaborate building at the back behind the front gate. The vertical lines expand to become the trees around the front of the picture that then push the eye out to make the larger vertical structures of the escarpments that form the background on the sides of the painting. The trees growing up the mountainside in the background are also vertical lines. In the middle of the painting there are a few trees and the two central pillars of the pagodas. Although these are clearly vertical structures, their blue horizontal rooflines create a striking balance in that portion of the painting.

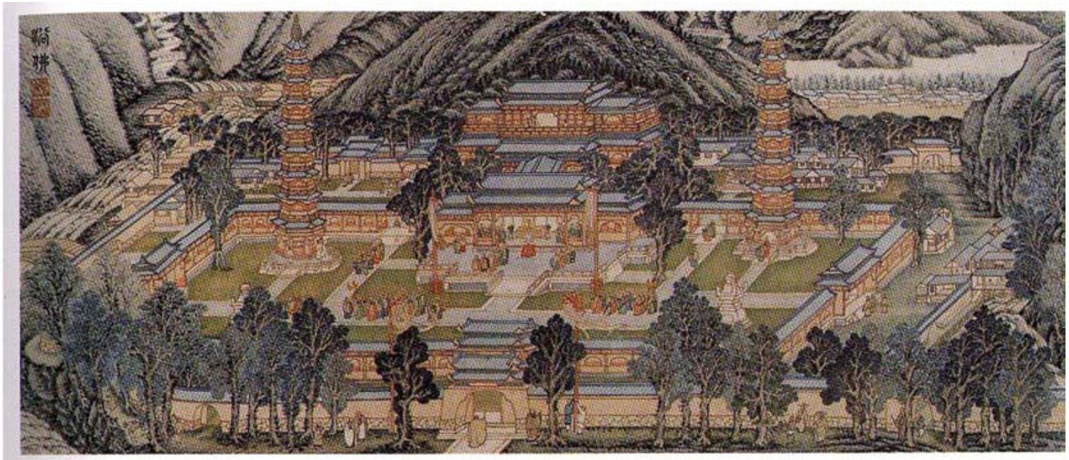


Figure 2. *Leaf 4, Washing the Buddha* from *Spring Morning in the Han Palace*

Repetition on a smaller scale may also take many forms in music. One type of repetition is called a motif -when a small, recognizable unit of a melody is used to create a theme. Two examples, the first by Native American flutist Carlos Nakai, and the second from a Beethoven string quartet, demonstrate how this kind of repetition can be used to build larger structures. Carlos Nakai's *Shaman's Call* begins with a motif--four different notes in a recognizable rhythm. After a pause, the motif is repeated, followed by another pause, and this time the motif is elongated. The result is a theme with two distinct sections: the first, two identical repetitions of the motif; the second, the motif elongated. Ludwig van Beethoven, in the first movement of *String Quartet #1*, presents a theme that has nearly the same structure as Carlos Nakai's theme. The only difference is the final note of the motif that Beethoven changes when he repeats it. The organization of Beethoven's theme, however, is identical to Carlos Nakai's: motif, motif repeated, motif elongated. Both of these examples show how a small-scale repetition of a motif can be used to create an entire theme.

Just as the artist used vertical and horizontal lines on a small scale and vertical and horizontal structures on a larger one, musicians can use repetition similarly. Carlos Nakai's *Shaman's Call* and Beethoven's *String Quartet #1*, use repeated and elongated motives to create the themes. On a larger scale, Carlos Nakai has repeated the whole theme, with its repeated and elongated motives to make up the first phrase of the piece. Beethoven's *String Quartet #1* uses nearly the same organization. Remember that when Beethoven repeats the motive, he changes the last note and then elongates it. In the larger-scale structure of the phrase, he makes a similar change on the repetition by making it much louder to fill out the sound. Both of these examples take a simple motif repetition used on a very small scale to make a theme, then and expand it on a larger scale to make a phrase.

Anticipation

Anticipation is a feeling of knowing what is about to happen, or a hopeful expectation of what is to come. For anticipation to work in art or music, the artist must manipulate how the audience will respond by preparing them for what is to follow.

In *Waiting for Guests by Candlelight*, the artist teases the viewer by showing only a portion of the interior of the house. The title of the work, and its images of guests waiting outside, and the host seated inside waiting for more guests to arrive, tell us that some kind of special event will be taking place. The viewer must anticipate and imagine what the interior of the house must look like, once we have stepped through the front door. This detail of the interior scene is not depicted, but one can anticipate, perhaps in great detail, how it might look and what might be happening inside.

Figure 3. *Waiting for Guests by Candlelight*

The cadenza and its conclusion from the first movement of *Brandenburg Concerto #5* by J. S. Bach is an enchanting example of musical anticipation. This work is a concerto grosso, a musical form in which a small group of soloists -- flute, violin, and harpsichord -- are accompanied by a slightly larger group of strings and organ. This movement has a very distinct theme that begins immediately played by the whole ensemble. This theme is repeated and developed throughout the movement by both the soloists and the full group. As the movement continues, however, the harpsichord becomes more and more prominent. (Some historians believe that harpsichord was prominently featured because Bach himself was playing the instrument in the 1717 performance.) Near the end of the movement is a weighty cadence followed by an extensive harpsichord cadenza that begins slowly and builds continuously for several minutes. By the end of this cadenza, even if one is hearing this work for the first time, one anticipates the return of the main theme. This theme, when it comes, is precisely what was anticipated, but the length and manner in which Bach has structured the cadenza before returning to the main theme, manipulates the listener into anticipating it. We might think of this as a musical version of the artist's teasing you into imagining what the interior of the house looks like.

Contrast

Contrast can provide variety, a change of pace, and a relief from repetitiveness. In the painting, *Hen and Chicks*, an anonymous thirteenth century scroll, contrast is used on several levels of both image and meaning. The contrast of dark and light colors is quite clear. The dark background causes the eye to be drawn directly to the light color of the hen. Additional interesting uses of contrast include size (large and small fowl), young and old (hen and baby chicks), and teacher and pupil (parent and children).

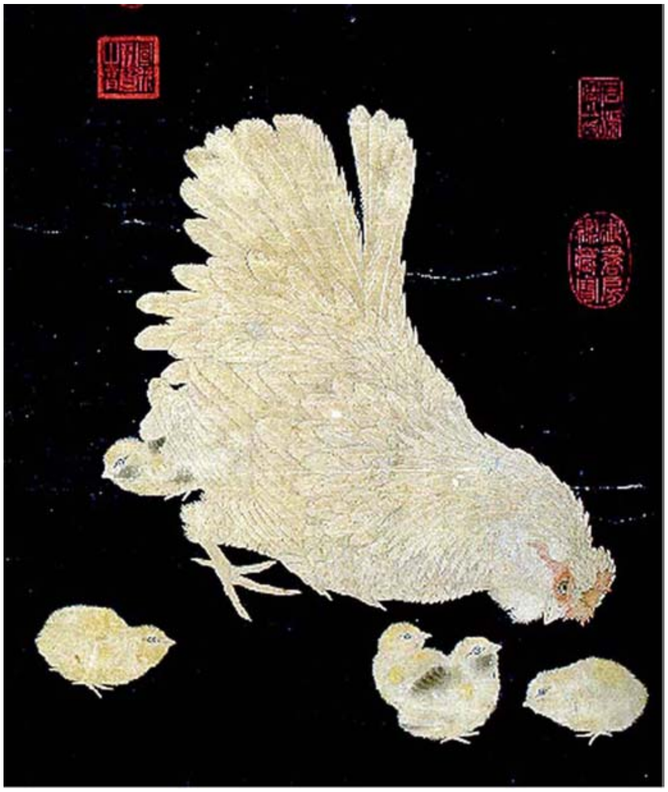


Figure 4. *Hen and Chicks*

Contrast is an important technique in a popular song by the British invasion band, The Who. *Pinball Wizard*, a song from the rock opera *Tommy*, contrasts several musical elements: *dynamic*s (degrees of loudness and softness), *timbre* (tone color or quality of sound), and *texture* (the number and importance of various musical lines). The tempo, with its basic quadruple rhythm, is constant with the opening basic pattern strummed on an acoustic guitar. Quite suddenly, a loud, clanging electric guitar chord interrupts this motoric pattern and completely changes both the dynamic and timbre of the song. When the voices enter, they sing the lyrics in unison -- a simple texture with instrumental accompaniment. This simple texture continues through to the last line of the verse, "Sure Plays a Mean Pinball," at which point the instrumental accompaniment drops out and the voices break into part harmony -- a more complex texture. In the course of fewer than 30 measures, the song contrasts dynamics, timbre, and texture.

Heightening of Tension

In both music and art, contrast may contribute to a heightening of tension, suspense, or some type of surprise. In another scene from *Spring Morning in the Han Palace,* the artist heightens tension in several ways. Most obvious is the tension on the bowstring, as the archer pulls it back. Additional, albeit more subtle, components also increase our sense of apprehension. The precipitous angle of the archer's body creates a sense of tension. He looks as if he's about to tip backward off the horse, whose front legs are at an odd angle, and it almost appears as if his left hand, the one holding the bow, is holding him up as well. Tension is also heightened by what we don't see: what is he aiming at with his arrow?

Figure 5. *Spring Morning in the Han Palace*

The last song from the album *Sergeant Pepper's Lonely Heart's Club Band* uses a typical device to heighten tension. The Beatles use a variety of creative elements in *Day in the Lif*e, and many critics believe this song represents the best of the group's compositions using this approach. Most of the Beatles' catalogue lists Lennon/McCartney as songwriters. With study and analysis, however, it becomes clearer that some songs are more "John" and others are more "Paul." *Day in the Life* is fascinating, because the first and last parts of the song are definitely by John. Even Paul said, "Those are John's bits." The middle section is a completely different song by Paul. During the process of recording, it was decided to use both songs and somehow fashion them together into a single work. The solution: create some kind of transition section between the first part of John's song and Paul's song. Then, after "Paul's song" a direct segue takes us back into the second part of John's song. This same transition device is then used to end *Day in the Life*. The transition consists of a 24-count constant underlying rhythm, during which symphony musicians were asked to begin at a quiet dynamic on the lowest note of their instruments, and then get increasingly louder and higher all the way up to the highest note of their instruments by the end of these 24 counts. They were instructed, however, not to do it together. The resulting sound becomes louder, higher and more raucous and culminates in a short silence. Then and only then, do we get the final chord of resolution -- just exactly the right chord that seems to go on forever, its great length helping to release the intensity of its build-up.

Suspense

Suspense is another technique to delay satisfaction. Heightening tension is often accomplished very aggressively -- perhaps as we say, "in your face." This increase of tension makes us feel more nervous, more anxious. To hold something in suspension defies gravity, holds us back from where we are going, as opposed to aggressively pushing us there. This detail, another scene from *Spring Morning in the Han Palace,* creates suspense in several ways. An important member of the family -- perhaps of the court -- is having her portrait painted, and a servant is nearby helping to keep the situation cool. The scene creates suspense by causing us to wonder "how long is this woman willing to sit still for her portrait?" Another servant is peaking round the corner -- or might she be an older wife, eavesdropping on the portrait session? The postures of all the principals indicate that it is time for something else to happen. How will this situation be resolved? Will the artist demand a few more moments to do the final touches on this particular part of the portrait? Or will the subject say, "I'm sorry, I've had enough of this for today and you'll have to continue later." The artist has created a suspenseful scene in which it is not yet clear who is actually in charge and how will this be resolved.

Figure 6. *Spring Morning in the Han Palace*

Beethoven's *Symphony #3*, the e-flat major symphony, subtitled the *Eroica* or *Heroic Symphony*, gives us an outstanding example of delay of satisfaction through suspense. An important part of the development section of any sonata form is what is called the retransition -- that section where we are returned to the home key preparing us for the first theme at the beginning of the recapitulation. In this example, Beethoven masterfully sets up this return by beginning with an increase of tension, but then suddenly backs away again -- an opposite approach to the "in your face" increase of tension used in the Beatles example we heard earlier. Beethoven, by holding back the tempo slightly as well as getting softer and teasing us with ever shorter bits of the first theme motif, creates suspense by forcing us to wait for the return of the whole theme.

Variation

Variation and its sibling, improvisation, are important devices for creating structure. In one sense, variation combines repetition with contrast. In order to recognize a variation of something, we must recognize enough of the original be aware that it has changed somewhat. The *Eleven-headed Avalokiteshvara* from seventeenth-century Tibet is a masterpiece of fine art and religious expression. This spectacular sculpture -- almost four feet high -- represents the most important Tibetan Buddhist deity, Avalokiteshvara. The exquisitely wrought details include the deity's one thousand arms, forming a halo around the figure. One thousand eyes appear on each palm. Many arms and palms, many heads, thousands of eyes -- each rendered slightly differently on the statue: variation.

Figure 7. *Eleven-headed Avalokiteshvara*

Improvisation, because it is usually based on some sort of framework, is a close relative of variation. A repeated rhythmic pattern called a "tala" is the framework for rhythmic improvisation in Northern Indian Ragas. The tala is always a specific number of beats, a cycle repeated until the conclusion of the performance. In this example, the tala is called *Jhaptaal* and it has a 10-beat pattern organized into two 5-beat groups of 2 +3 -- 2+3. The close connection between variation and improvisation can easily be heard because the repeating-- 1-2, 1-2-3, 1-2, 1-2-3, -- allows the percussionist freedom to be creative but still within the specified pattern. The tabla performer in this excerpt is Alla Rahka, considered to be one of the most famous tabla performers of the twentieth century. During this performance, the sitarist, Ravi Shaker, is sitting next to Alla Rahka and counts out the *Jhaptaal* pattern using traditional Indian hand signs.

Surprise

Surprise is something that catches us unawares. It can be a sudden realization, a coming together that causes us to smile, to feel delighted, or it might be shocking, something that is totally unexpected that blindsides us. Both manifestations are common and necessary techniques used in art and music to manipulate the ebb and flow of a work.

Delight

This very famous painting is from a scroll called *Winter Play*, and what a beautiful example of the surprise of delight it is. The delight here is on two levels. The first is the subject of the painting itself: the children, playing in the enclosed courtyard, are surprised and delighted with the kitten. The second is our response, remembering how happy children can be by seeing how the artist has captured this delight through the children's facial expressions and postures.

Figure 8. *Winter Play*

A song by the Beatles creates an unexpected, but delightful surprise through the transition to the final song in a medley -- a group of several songs all interconnected. The aggressive beat and "heavy metal " guitar solo of the preceding song suddenly drops out, and we're left with rhythmic piano chords in a four-beat feeling that delightfully transforms itself into *The End*, a song with beautifully simple lyrics. Lyrics that are not pushy, not aggressive, but very, very soothing: "And in the end, the love you take is equal to the love you make." With the words "And in the end, the love you take" the meter continues in four . . . but at the words, "is equal," the meter changes very subtly to a three-beat feeling. The piano chords are played at the same speed, but now they are grouped in patterns of three instead of four. This transformation is not expected at all. It's very surprising, the turn that it takes -- the change of meter, the change of style, the change of tonal center -- surprising but delightful nonetheless - it brings a smile.

Shock

There is also Surprise that is shocking -- the kind that elicits from us a response of "What was that?!" "Where did that come from!?" The *Left-Sided Angel*, a statue near the main library on the campus of Iowa State University has elicited that sort of response from viewers for many years. We don't expect an angel to have only one wing, one leg, one arm, or appear so battle weary. The posture of the foot also makes difficult to determine if the angel is descending from heaven or ascending to it. For some students, their shock goes no further than rejection, but for many others, it results in ironic humor. Recently a student walking past muttered, "My parents said college would cost an arm and a leg."

Figure 9. *Left-Sided Angel*

Igor Stravinsky's *Rite of Spring* produced riots at its Paris premiere in 1913, although most of the rioting took place before the piece was ever performed. The kind of surprise we feel in one particular passage is definitely shocking. Stravinsky sets us up by using a repetitive phrase structure that is altered very subtly by changing some of the rhythmic figures and instrumentation, but always softly. These call-and-response phrases are repeated so consistently that we anticipate they will continue to a satisfactory close. Suddenly, and for no apparent reason, the pattern is shattered by a booming bass drum and tam-tam. Immediately the dynamic level is ratcheted up dramatically with raucous chords and trombone smears. It is unexpected, and the surprise is shock, not delight. This passage is still shocks many listeners, even after repeated hearings!

Conclusion

As a conclusion to the series of lectures, I show a six-minute segment from the last movement of Beethoven's *Ninth Symphony* that incorporates the elements discussed in those lectures. Included are the devices and techniques of variation, repetition, and contrast, along with human responses of delay of satisfaction, heightening of tension, surprise, shock, and delight. This presentation is intended to work on its own without commentary. The descriptions in words with appropriate art examples are coordinated on the slide show to correspond to the appropriate technique in the music.

Student reactions to this lecture series have been very positive. A number of students ask either after class or via e-mails how to find the works of art on campus. Many comment that including popular music examples to demonstrate these formal techniques, causes them to listen to these songs differently. Additionally, we discovered that the "tag-team" approach to delivering the lectures -- an expert from each of the areas alternating music and art, technique and response -- helped students better understand how universal these formal elements are.

Web-Based Instruction and Diversity Issues: Planning for Inclusiveness A Literature Review and Practical Guide for Educators

Tracey Stuckey-Mickell
Northern Illinois University

Abstract

In efforts to create and maintain an optimal web-based learning environment for all learners, researchers are addressing diversity issues within web-based instruction (WBI). This work reviews and synthesizes current literature, applying approaches in creating and facilitating inclusive courses to web-based instruction resulting in a synthesis of ideas and a practical guide for educators interested in designing inclusive WBI. Appendices referred to in the abstract will be provided for participants at the convention.

Background and Overview

Web-based instruction (WBI) is becoming widely used in educational institutions across the world. Learners of different cultures, ages, genders, classes, and academic level are introduced to the world through online learning. As we scurry about in our attempts to optimize the online learning environment and maximize student learning, we must consider how issues of diversity, such as culture, gender, and class, interact with our current concerns about online learning. In order to reach as many learners as possible, we must consider how to design online instruction that is sensitive to the needs of a diverse student population.

Diversity issues in WBI are fairly new in the research literature. Many of the studies reviewed have been published within the last two to five years. In order for educators to make the best use of this research, integrate its findings into their work and further their research, it is important to review current research and thought, synthesize findings and develop a collective pool of theoretical assertions, supportive research results and strategies or interventions. This work is a discussion and compilation of current research findings and theory on design/facilitation issues and inclusiveness in education and global WBI. These findings are synthesized into a practical guide of issues to consider for educators who want to design web-based instruction that is sensitive to diverse groups of learners.

Issues in Design and Facilitation

Educators must remember that WBI is not simply face-to-face instructional activities transferred to the online environment. In the optimal online learning environment, both the role of the tutor/instructor and learner are transformed. Learners are given more responsibility for their own learning, while tutors are pushed into a support role. Designing with a variety of approaches and learning resources can create a more inclusive environment. It is important to remember that well-designed WBI considers the diversity of learners begins at the planning stage continues throughout administration. The following studies suggest points to consider when designing and administering WBI.

Shaw (2001) gives us some considerations for designing online instruction by considering several different approaches to the design of online course materials. He describes the online course design process as "new and complex" where designers must make decisions on a variety of resources, activities and online learning features (e.g, asynchronous, synchronous, websites, virtual classrooms, application-sharing), as well as content. He reviewed five different approaches to designing the online learning experience. Each approach was evaluated. He concludes that approaches that fully exploited the features of online course delivery gave learners the most control and tutor availability, but also cost more and required more time and effort to develop and facilitate. Providing this environment requires a team collaboration of designers, content experts, and technological experts. Educators interested in developing WBI should consider the resource commitment involved to create optimal WBI.

Levin, Levin & Waddoups (1999) investigated how multiplicity can enhance the online learning experience. Through analysis of survey and interview data gathered from the participants of three professional teacher training courses in technology integration, they found that the employment of multiple contexts for learning, multiple instructional media, multiple instructional formats, multiple learning activities and multiple assessment techniques played a key role in enhancing the online learning experience. Table 1 summarizes their findings. The authors assert that using multiple resources and methods in WBI expose students to various ways to use technology beyond the course and increased learner engagement. Though front-end investment is high, the possibilities for a more powerful learning environment in increased in the long run.

Multiple Learning Contexts	Multiple Instructional Media	Multiple Instructional Formats	Multiple Learning Activities	Multiple Assessment Techniques
Online class	Email	Project groups	Learner surveys	Peer assessment (other course participants)
Student work environment (authentic application)	Internet mailing lists	Reading discussion groups	Web development	Tutor assessment
Simulations	Web archives of list messages	Real-time interactions	E-portfolio	Self assessment
Informal learning groups	Faculty & student web pages	Simulations	Technology integration proposals & plans	Assessment by outside audiences (educators, colleagues)
	Asynchronous conferencing & synchronous chat (Web Board)	Whole-class student presentations	Technology integrated lessons	
	Streamed audio/video (RealPlayer)	Electronic field trips	Resource development	
	Web-based database system (used to submit, display and assess work)	Online readings and textbooks	Presentations	
	Telephone conferences			

Table1. *Summary of Multiple Elements Used in the Design of an Online Course (compiled from Levin, Levin & Waddoups, 1999)*

Murphy and Cifuentes (2001) used a student-centered approach to investigate design elements of online learning environments and how they relate to the different ways in which students learn. After conducting and online course and using qualitative methods of inquiry to discover student reactions to the online learning process, Murphy and Cifuentes concluded that learners worked with the e-moderator to familiarize themselves with each other, establish goals and objectives for activities, develop timelines and task lists in groups, and used both synchronous and asynchronous communications as well as face-to-face group meetings to facilitate collaboration. Learners also used peer and tutor feedback to assess their progress.

There are two key points from this study. First, dialogue is at the center of the quality online learning experience. Second, in the online learning environment, structure is also important. Both of these elements serve to keep student frustration levels low and student engagement high. The authors assert that when designing an online course, the designer should strive to strike a balance between both dialogue and structure. Design should give way to the development of an online community of learners where collaboration is encouraged and maintained throughout the course.

Once the course has been designed and the technological environment implemented, the facilitator is ready to guide students in their online learning experience. Experts agree that the role of the instructor is substantially different in the online learning environment as opposed to the face-to-face environment. Additionally, students find substantial differences in the group dynamics when interacting with others online. In order to facilitate the online course successfully, instructors (or e-moderators) must recognize the role change and be aware of the group dynamics of the online environment.

Salmon (2001) presents a model of online teaching and learning that provides insight into the roles of both the learner and the e-moderator. The model is a five-stage progression that outlines student and e-moderator participation and interaction in terms of required learner technological skills, e-moderator skills, and level of learner interaction throughout the online course (see Figure 1). This model gives the e-moderator a framework by which to guide learners through their experience in the online learning environment (see Table 2). For example, in stage 1, Access and Motivation, learners are faced with initially accessing the online environment and becoming motivated to log on. For the e-moderator, providing technical support and encouraging learners is important at this stage.

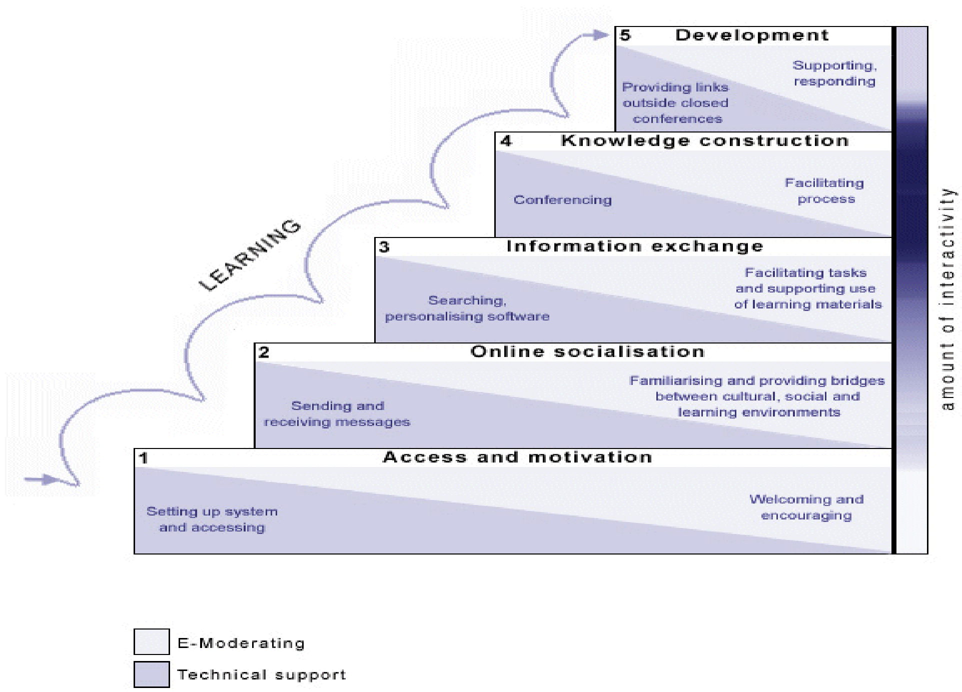


Figure 1. Gilly Salmon's 5-step model of online learning: learner challenges and e-moderator strategies 2001. [Online] Available: http://oubs.open.ac.uk/e-moderating/fivestep.htm

	Stage 1 ~ Access & Motivation	Stage 2~ Online Socialization	Stage 3~ Information Exchange	Stage 4~ Knowledge Construction	Stage 5~ Development
Learner Tasks & Challenges	Setting up and accessing system	Sending & receiving messages	Searching & personalizing software	Conferencing	Providing links outside close conferences
E-moderating Skills	Welcoming & encouraging	Familiarizing & providing bridges between cultural, social & learning environments	Facilitating tasks an supporting use of learning materials	Facilitating process	Supporting & responding
Interaction Level	Low	Medium	High	High	Medium

Table 2. Salmon's Stages

A key point to consider about Salmon's model is that certain skills are important at different times for both learners and e-moderators as they move through the online course. Additionally, e-moderators should expect different levels of learner interaction as users become oriented to online learning and begin to develop strategies for meeting their own learning needs. Educators should also consider how individual differences might impact the learners' experiences at each stage.

Varvel (2001) reviews literature on "traits" of the successful online learner and discusses the need for online instructors to assist learners who may not possess these traits. He advocates for nurturing those traits in order to reach every student in the online learning environment. He reviews categories of problems faced by the "non-ideal" online student and some helpful techniques within each category that e-moderators can use to "facilitate every student" in an online course (see Appendix 1). A key point to consider is that, similar to the face-to-face classroom, online learners have different skill and comfort levels to be addressed. In order to lower student attrition, e-moderators should be aware of the different challenges and have some plan to address them.

In considering the approaches to and concerns about online learning design and administration, educators must also consider approaches and concerns to reaching students across culture and gender. The online classroom is a global classroom, where students from many different walks of life can participate in learning experiences together. E-moderators must consider some basic premises and approaches to facilitating a culturally diverse population of students. Many experts believe that the process begins with an exploration of self.

Diversity Considerations in Self and Others

Merriam, Johnson-Bailey, Lee, Kee, Ntseane and Muhamad (2001) discuss positionality with regard to conducting research. Each author conducted research with both same-culture and different-culture participants and discuss the "insider-outsider" phenomenon. They conclude that culture serves to socialize and shape the positionality of individuals, their internal power relationship to others and subsequently their representation within the "insider-outsider" phenomenon. Though the thrust of their work centered on implications for conducting research, there are implications for educators to consider.

Effective instructors must have some idea of how his or her own background, biases and experiences (or positionality) impacts his or her philosophies and practices in the classroom. Positionality becomes increasingly important when or facilitating diverse student populations in an online learning environment. Since face-to-face interaction may be limited or nonexistent, one must take care in the messages one sends through course design, selection of materials and course facilitation.

Whether or not we realize it, positionality plays an inherent role in how we design our courses (e.g., the approaches we use and type and amount of activities we select), our selection of materials (e.g., what content is presented as important or unimportant) and the way we facilitate the course and interact with students (e.g., our communication style, our perception of students and our ways of relating information). Students' perception of the instructor and the course may be colored by the instructor's positionality. Whether or not the instructor is viewed as a concerned and helpful guide to learning or an indifferent gatekeeper of intellectual knowledge depends upon how the instructor is represented through the course. Since online learning may involve little or no face-to-face interaction, it is increasingly important for e-moderators to reflect upon their own positionality and how it may manifest itself to the learners in the online course.

Coleman (2002) discusses how addressing issues of diversity creates opportunities for instructors to examine their own beliefs, views, and attitudes toward different groups in preparation for and during the facilitation of their courses. In fact, this type of self-development and self-acceptance, she emphasizes, influences instructor classroom performance in how differing perspectives and dealings with diverse groups are handled. She asserts that our comfort level with "differentness" in our worldview is dependent upon our comfort level with ourselves.

Coleman advocates for the examination of one's own biases, discomforts and stereotypes about other groups. She also describes helpful tips for educators (see Appendix 2) to use to begin examining their beliefs and attitudes as well as teaching tips to employ in the classroom. Key points concern the relationship between self-understanding and acceptance to the handling of diversity in the classroom. Simply claiming to "not see color" in your students does not lead to equitable treatment in the classroom. One must examine and accept his or her biases and develop conscious methods of dealing with them in order to work well with diverse populations in the classroom. This is true also for dealing with global online learning. E-moderators must examine their beliefs about individual differences in order to address biases, value differences, and stereotypes in order to deal effectively with a diverse population of learners.

Zuga (1999) reviews ways to address women's ways of knowing and technology education through a discussion of feminist theory as it relates to critically restructuring technology education to address women's ways of knowing. Though not all online courses are technology education courses, Zuga's idea of addressing women's ways of knowing is greatly applicable to any online course content.

Through a discussion and analysis of research on cultural differences for women, masculine bias, and moral development of woman, Zuga asserts that women may indeed think and learn quite differently from men and that these differences should be addressed in the classroom. She also points out some possible ways that educators can begin to deal with these differences in technological learning environments (see Appendix 3). Some online courses may not have students from different countries or cultures, but most if not all, will include female students.

Of unique importance are the issues faced by many working mothers as they embark upon online education. Santovec (2002) describes online learning as the "third shift" for many women who juggle work and children while furthering their education. Often many of them feel that online courses are the only way to continue taking courses. In the discussion of results from an AAUW study on women and online learning, Santovec urges designers of online education to consider challenges that female online students may face and involve women in the course design process. There is also a need for further research on how women who have families and jobs outside of the home navigate the online learning environment.

Curriculum Integration and Transformation

In order to create courses that are truly inclusive across culture and gender, certain changes to traditional curriculum must be made. The process of revising curriculum to achieve this goal is commonly referred to as curriculum integration or curriculum transformation.

Scott (n.d.), in a review of relevant literature, discusses methods for integrating race, class, gender, and sexual orientation into college curricula. She describes curricular integration as "a dialectical process between course restructuring or transformation and changing classroom environments." She insists that both the course and learning environment must change to reflect true integration. She cites Higginbotham (1988) listing three components essential to integrating diverse populations into course curricula:

1. broaden student personal knowledge on diversity issues;
2. transform content and course structure to meet diverse student needs; and
3. adjust dynamics of the classroom to provide a safe environment that is inclusive of all students.

Through these revisions, a course is transformed into a process in which personal truths are investigated and discussed whereby students are able to question biases and stereotypes while reflecting and learning about course content. As she moves from theory to application, Scott presents useful and practical considerations and methods for

Banks (2002) presents another approach to curriculum revision. Banks discusses curricular reform as a transformation process existing on a continuum whereby educators use different approaches that reflect varying levels of integration and reform (see Appendix 3). The most desirable approach, Social Action, is characterized by activities and projects that allow students to make decisions and take action regarding the problems, concepts, and issues covered in the course. In this approach, students are encouraged to move past learning about shallow social indicators of a particular group. Instead they are challenged to question traditional thought, create new awareness, synthesize ideas, make personal decisions, and take action.

Both approaches can be integrated into WBI through careful analysis of the learners, course objectives, and facilitator self-examination. Educators, through learning about self and others, can create an environment that fosters understanding of and across individual differences within any course content.

Creating Bridges through Web-based Instruction

Through online course delivery, educators can maximize educational efforts and reach populations who may otherwise be marginalized. It is of considerable importance that the broadening spectrum requires educators to increase their understanding of cultural differences and revise courses to be inclusive. Liang & McQueen (1999) identified student expectations, facilitator roles, grouping preferences, and communication styles as serious concerns that may impact the online learning environment across cultures. In spite of the seriousness of these issues, there is significant but limited research on WBI and multicultural issues. Educators are only beginning to examine how online global learning, gender and multicultural issues intersect.

The unique nature of the group dynamics and interaction in online learning environments have prompted educators to examine the online classroom as a learning environment that is conducive to the discussion of diversity and cultural differences across geography and culture (Cifuentes & Murphy, 2000; Chabon, Cain, & Lee-Wilkerson, 2001). The ability to combine courses from different geographical and demographical campuses provided for more diversity in the learning community along with usual benefits of online learning. However, there were some challenges. For example, the misinterpretation of the written word was one challenge. (Chabon, Cain & Lee-Wilkerson, 2001). The authors advocate for solid ground rules regarding interactions and reflection upon statements posted due to the lack of face-to-face interaction and the high sensitivity of the topic.

Conclusions: Creating and Maintaining Inclusive WBI

The major purpose of looking at online learning environments and multicultural education paradigms is to discover how we can meet the needs of the diverse student population we are able to reach with the online classroom. Many of the benefits of online learning, such as “anytime, anywhere learning” and student control of learning, may have embedded problems such as the loss of nonverbal communications, student perception about “impersonal” learning communities and increased time requirements on the part of both learner and facilitator.

Much of the concern centers on awareness. E-moderators and online course designers must be aware of diversity in learner needs and adjust design to be sensitive to those needs. Language issues, cultural communication differences, positionality and attitude, multicultural course content and structure, along with ease of use with regard to technology are all issues in WBI. Technology has greatly improved our ability to span culture and distance. If educators aim to design a learning environment that is optimal for all learners, these issues must be considered during the design and facilitation of WBI.

References

Banks, J. A. (2002). Curriculum transformation (Chapter 3-pp. 21-34) An Introduction to Multicultural Education, 3rd Ed. Boston: Allyn & Bacon.

Chabon, S. S., Cain, R. E., & Lee-Wilkerson, D. (2001). Facilitating those dreaded discussions on diversity, through threaded discussions: An inter-institutional, internet-based model. Distance Edcuation, 22 (1), 137-143.

Cifuentes, L. & Murphy, K. (2000). Promoting multicultural understanding and positive self-concept through a distance learning community: Cultural Connections. Educational Technology Research and Development, 48 (1), 69-83.

Coleman, L. The influence of attitudes, feelings and behavior toward diversity on teaching and learning. [Online]. Available: http://www.colorado.edu/ftep/support/diversity/div05.html Colorado: Faculty Teaching Excellence Program. Accessed: 4/18/2002.

Levin, J. & Levin, S. R. (1999). Multiplicity in learning and teaching: A framework for developing innovative online education. Journal of Research on Computing in Education, 32 (2), 256-267.

Liang, A. & McQueen, R. J. (1999). Computer assisted adult interactive learning in a multi-cultural environment. Adult Learning, 11 (1), 26-29.

Merriam, S., Johnson-Bailey, J., Lee, M., Kee, Y, Ntseane, G. & Muhamad, M. (2001). Power and positionality: negotiating insider/outsider status within and across cultures. International Journal of Lifelong Education, 20 (5), 405-416.

Murphy, K. L. & Cifuentes, L. (2001). Using web tools, collaborating, and learning online. Distance Education, 22 (2), 285-305.

Salmon, G. (2001). Psychological and group learning perspectives: Their relevance to e-learning. Presented at E-Learning, London, February, 2001.

Santovec, M. L. (2002), Balancing act: online learning becomes the "third shift" for women. Distance Education Report, 6 (2), 1-2.

Scott, B. M. (n. d.). Integrating race, class, gender and sexual orientation into the college curriculum. Unpublished Manuscript.

Shaw, K. (2001). Designing online learning opportunities, orchestrating experiences and managing learning. In Stepenson, John. (Ed.), Teaching and Learning Online: Pedagogies for New Technologies. London: Kogan Page.

Varvel, Virgil E. Jr. (2001). Facilitating Every Student in an Online Course [Online]. Available: http://www.ion.illinois.edu/pointers/printerfriendly/ionpointers0301Print.html
Illinois Online Network Website: http://illinois.online.uillinois.edu/ Accessed: 4/18/2002.

Zuga, K. (1999). Addressing women's ways of knowing to improve the technology education environment for all students. [Online]. Journal of Technology Education, 10 (2). Available: http://scholar.lib.vt.edu/ejournals/JTE/v10n2/zuga.html Accessed: 4/18/2002

Appendices

Appendix 1

Varvel's categories, problems and techniques for "facilitating all students." (compiled from Varvel, 2001)

Time Management Skills	Discipline & Motivation	Synergy and the Online Learning Community	*Communication Skills2	Technophobia	*Access4
Lesson Length: Small information chunks	Motivate beyond the grade: Use discovery, feedback	Provide a course philosophy: Part of course info, model it throughout	Be a model: Stay positive, remove inappropriate posts, provide a unifying voice, clarify or extend	Pre-course orientation: Arrange a face-to-face first meeting or online tutorial.	Disability Accessibility: Be aware of accessibility concerns.
Take the Lesson Along: Make online pages printable	Make your presence felt: Responding to posts; higher level thinking; instructor's page (picture, short biography.	Structure discussion into course: Build into design; guidelines (e.g., length, content, number, etc.).	Give instructions: Clear & concise; guidelines (length, content, etc.)	Support Services: Before beginnimg, instructions for tech assistance.	Network access & bandwidth: Design for slowest network connection, 28.8 kbs modem; alternate versions of high bandwidth content (e.g., CDROM). Limit use of graphics.
Effective Syllabus: Full explanations	Direct questions: Question lagging students; "call on" specific students to answer questions	Structure the discussions: Organize topics (separate forum for each), lead with first post.	Provide for communication in the course philosophy & requirements:	Patience: Tehnophobic students, help them keep up.	
Make suggestions: Tips on managing time	Avoid interruptions: Suggestions to help with deterrents.	Be engaging: Use humor, share anecdotes & experiences	Provide motivation & encouragement: Send individual messages		
Provide clear posting requirements		Break the ice: Use an icebreaking activity at beginning	Remedial efforts: Be aware and understanding about communication difficulties		
		*Fix problems promptly1: Inappropriate communications addressed immediately.	*Understand forms of online talk3: Use forms to your benefit, provide students with resources		
* See next page for detailed notes			Meanings within meanings: Be able to recognize subliminal meaning (e.g., emotional stress or hostility) in student posts & address accordingly.		

Appendix 1 (continued)

Detailed Notes: Varvel's Categories

1Some examples of disruptive students are:

The Know-It-All – self explanatory, right??
The Mutineer – chronic complainer
The Lagging Belligerent – falls behind and it visibly angry about it
The Attacking Belligerent – similar to the Mutineer
The Staller – chronically late student
The Non-Participant – also known as "lurker"

For other examples, see Varvel (2001) [Available Online].

2There are two primary modes of communication online: asynchronous (ex: conferences, email) or synchronous (ex: live chat). Varvel provides tips for both:

Asynchronous

- Start major topics with an explanatory post.
- Narrow topics to smallest units to reduce clutter in discussion forum posts.
- Organize forum to correspond to course flow.
- Be aware of cultural patterns in the manner in which students post.
- Respond frequently, but save "nice job" posts for individual student conferences or email.

Synchronous

- Limit group sizes (4-12 is optimum).
- Employ "crowd control" methods, if necessary (ex: Students take turns, agree not to interrupt, etc.).
- If possible, use audio for instructor feed.
- Allow some socializing before and after.
- Be prepared to stretch sessions out over time, possibly with different times for different groups or topics to address student availability.
- Post an agenda to keep focused and to give students time to prepare.
- Use names to preface individual responses.
- Always have a backup plan. Technology isn't foolproof. Prepare a different chat engine system or another time/date.

3There are several forms of online talk. Varvel lists some of them:
Control Talk-providing guidelines & instructions
Humor-entertaining & enlightening spirits
Acronyms-short meaningful acronyms (IMHO-in my humble opinion; RFLMAO-rolling on the floor laughing my behind off; BRB-be right back)
Emoticons-text queues to emotional states :o) :> :o(
Androgical Approach-remove gender from language unless appropriate

4 Website on accessibility:
http://www.apple.com/disability/

Also, contact your institutions support services offices for assistance.

Appendix 2
Tips for Examining Feelings, Attitudes & Belief Toward Diversity. Adapted from Coleman, (2002)

Thought questions for examining self-behavior
Attitudes and Beliefs

1. How do I feel about women and students of color in academia?
2. Can I see beyond a student's gender, ethnicity or racial background and see an intelligent, evolving individual?
3. Do I have different expectations for students based upon gender, ethnicity or racial background?
4. Am I surprised when female, Black or Latino students do well?
5. Do you see ability and intelligence as limited capacity or an expandable commodity?

Treatment of Students

1. Have I encouraged bright students of color and women to assist me with research or join my research team/lab?
2. Have I encouraged bright students of color and women to pursue graduate school?
3. Do I encourage students of color and women to pursue their interests?
4. Do I treat students differently based on my feelings about their gender, ethnicity or race?

Specific Teaching Tips

1. Advise every student to take an ethnic studies or women's studies class regardless of ethnic background or gender.
2. Attempt to provide a variety of perspectives on all topics you teach (e.g., Eastern view, marginalized population view, etc.)*
3. Attempt to help students understand how we (as students and individuals) construct meanings attached to different groups (e.g., ability, racial, ethnic, gender, social class, etc.).*
4. Ask students to write a paper about the meaning of race (and its relevance for your academic discipline). OR Ask them to speak with a person of different race or ethnicity about what it means to them and write about it.*
5. Do not allow students to make unsubstantiated statements or sweeping generalizations about different groups of people without requiring them to conduct research to support such claims.***
6. Try to answer personal questions about gender and racial issues as honestly as you can.*
7. Do not avoid controversy. Have students talk openly about stereotypes, where they originate and why/how they are perpetuated.*
8. Encourage students to engage in individuation. This is the process by which we attempt to move beyond stereotyping a person to seeing him or her as an individual.*
9. Stop the class when students are having trouble talking about a topic. Have them discuss their discomfort, especially their feelings.**
10. Do not allow students to attack each other in class. Have them use "I" statements.***
11. Encourage students, especially those who are angry about issues to discuss with you individually in greater detail.**
12. Assist students in understanding how they specific racial or gender makeup may help them to get more in touch with their humanity. *

*These are easily accomplished by the use of a special forum in the online classroom.
**May warrant a special email or other individual means of contact in the online classroom.
***Can be presented in the course information guidelines or groundrules.

Appendix 3

Addressing Women's Ways of Knowing in the Classroom (Adapted from Zuga, 1999)

1. Establish contexts for ideas.
2. Revise language and actions to be inclusive.
3. Use complete and thorough explanations for ideas.
4. Incorporate students' values into the curriculum.
5. Create activities/assignments with women's ways of knowing, values and interests in mind.
6. Study feminist texts to increase awareness about ways of knowing and student attitudes/beliefs about technology.

For Further Reading

Bame, E. A., & Dugger, W. E. (1990). Pupils' attitudes and concepts of technology. The Technology Teacher, 49(8), 10-11.

Belenky, M.F., McVicker Clinchy, B., Rule Golberger, N., & Mattuck Tarule, J. (1986). Women's ways of knowing: The development of self, voice, and mind. New York: Basic Books.

Donovan, J. (1994). Feminist Theory: Intellectual Traditions of American feminism. New York: Continuum.

Hooks, b. (1994). Teaching to Transgress: Education as the Practice of Freedom. New York: Routledge.

Miller, J. L. (1993). Constructions of curriculum and gender. In S. Knopp Biklen & D. Pollards (Eds.) Gender and Education. Chicago: University of Chicago, pp. 43-63.

Approaches to Multicultural Curriculum Reform (Banks, 2002, p. 30)

Level 4: Social Action
Students make decisions on social issues and take actions toward solving them.

Level 3: Transformation
Curriculum structure is changed to allow students to view concept, issues, events, & themes from the perspective of diverse groups.

Level 2: Additive
Content, concepts, themes, and perspectives are added to the curriculum without changing its structure.

Level 1: Contributions
Focuses on heroes, holidays, and discrete cultural elements.

Performance Systems Analysis in a Major Hospital in the Southeast: *A Case Study on Patient Education*

Sookyung Suh
Marc Knowles
E Shen
Rebecca Sallisky
Florida State University

Abstract

This paper will discuss the project that was spawned from a course at Florida State University entitled Performance Systems Analysis (PSA). This PSA course allows students a chance to work with clients who have identified performance that does not meet organization expectations. The detailed procedures of the 4 month-long analysis will be discussed along with the methodologies that were used. The recommended solutions will be identified as well as evaluation. Students and professors who are interested in learning about PSA will benefit from this presentation.

Introduction

This project was spawned from a course at Florida State University entitled Performance Systems Analysis (PSA) in the College of Education, Educational Psychology and Learning Systems Department. Our team in PSA took the role of a performance improvement consulting group for a major hospital in southeastern U.S. Through a detailed performance systems analysis, including a gap and cause analysis, our team recommended solutions to close the identified gap through an appropriate intervention. The stakeholders decided to implement most of the recommended solutions that our team suggested. After implementation, our PSA team evaluated the performance improvement efforts.

The services of our team were secured by the Learning Systems department of this major southeastern hospital in September, 2002, to address concerns of patient education. The patient education manager, whose department is a division of Learning Systems, was concerned that the education of patients by clinical staff members was not occurring at an adequate level. After an initial interview with the client, our team proposed data collection methods to determine the gap between desired organizational and workforce performance and identify the causes of these gaps. The causes of these gaps were identified using Gilbert's (1978) Behavior Engineering Model (BEM). The causes were then prioritized using a prototype electronic performance support system (EPSS) for performance systems analysis developed by Learning Systems Institute, Florida State University, for the United States Navy, in which each cause was rated on the levels of magnitude, cost, and urgency. After identifying and prioritizing the causes of the found gap, we identified solutions to close the gap by systematically analyzing the causes. Our team again used the EPSS, which contains a possible solution generating worksheet.

With regards to the hypothesized performance gap, patient education is for the good of all involved in medical procedures. It takes into account the needs of the patient while in the hospital, upon release, and coping with their ailments in general. There are many domains the patients must be educated on including pain management, self-medication, treatment options, following up with doctors, and coping with ailments, to name a few. Patient education must occur and be documented for several reasons. The most obvious is to gain accreditation from the Joint Commission for Accreditation of Healthcare Organizations (JCAHO). During a past paid consultation from JCAHO, it was noted that pain management was not being taught/documented at an adequate level and this needed to be improved before JCAHO returned; consequences on the subsequent visit could be detrimental to the hospital if rectifying actions were not taken.

Data Collection

Organizational Analysis

Learning Systems Department is a division of the hospital in which we performed our analysis. Learning Systems is adjacent to the hospital, and delivers internal education services. Patient education is a major part of these services. This service is of great importance because it gives essential information to patients regarding how to cope with illnesses, how to medicate themselves, and many other domains made standard by the Joint Commission on Accreditation of Healthcare Organizations. Learning Systems' mission is to deliver adequate patient education to 100% of patients, during their stay at the hospital.

Clinical staff members at the hospital are supposed to serve as patient educators regardless of the realm of their work. This means that nurses, radiology technicians, pharmacists and others are expected to use "teachable moments" to convey important information to patients and their families. Even though all clinical staff members are responsible to deliver patient education, nurses have the most impact because they interact with the patients in the setting most frequently and for the longest period of time.

A strategy that is currently in place to ensure patient education is known by some staff is a competency check off class. Unit coordinators for each floor are required to learn about patient education and complete a competency check off class with a specialist nurse.

The internal stakeholders were identified as the patient education manager, nurses, and patients while the external stakeholder was the Joint Commission on Accreditation of Healthcare Organizations

Performance Analysis

Gap Analysis To conduct a gap analysis, desired work performance and actual work performance data were collected. To collect actual work performance, a nurse questionnaire was created and individual interviews were conducted. The first questionnaire was designed to find out the mindset of nurses toward patient education, and whether the nurses educate patients in an appropriate way. After analyzing the first questionnaire, it was discovered that nurses do educate patients but they do not document it due to a variety of reasons.

Desired work performance of Organizational level

The measurable ***Organizational level*** desired goals were identified as:

1) 100% satisfaction rating by the "Patient Satisfaction Survey" in regards to "Patient Education". There are six questions that relate to patient education on this survey:
    a. I am able to care for myself (in-patient survey)
    b. Nurses knowledgeable about conditions (in-patient survey)
    c. Nurses responsive to my pain (in-patient survey)
    d. Received adequate information before leaving (in-patient survey)
    e. Staff was concerned about my comfort (out-patient survey)
    f. Told me what to expect during experience (out-patient survey)
2) Zero monthly complaints due to unsatisfactory education by nurses to patients in all units at the major hospital in the southeastern U.S.

Desired work performance of Work level

3) Nurses will be able to perform patient education at the appropriate time.
   Nurses should have the skills to apply the following four steps in performing patient education with 100% of the patients, and documenting it along the way.
    - Assess: Define patient and family needs and concerns; observe readiness to learn. Evaluate your assessment throughout the process.
    - Plan: Set learning objectives with the patient; select materials. In planning, nurses should be able to make a road map that guides patients to their objectives.
    - Implement: Put the plan in motion; help patients reach mutually agreed objectives. This is when the nurses conduct most of their education of patients.
    - Evaluate: Determine if the planned educational tactics were effective in delivering knowledge.
4) Documentation: Nurse will be able to create a written history of the patient education. All the records should be able to be read by other health care team members and a JCAHO review board.
5) Nurses should regard patient education as one of their primary tasks. According to a survey of nurses regarding "Patient Education", 95% percent of the returned answers will mark "Patient Education" as important as their traditional primary nursing tasks.
6) Patients will be 100% educated concerning what patient education is and what their rights as a patient are with regards to patient education.

Actual work performance of Organizational level:

1) The actual satisfaction rate by the "Patient Satisfaction Survey" in regards to "Patient Education" is shown by the following six questions from the survey related to patient education:
    a) I am able to care for myself (in-patient) 93%
    b) Nurses knowledgeable about conditions (in-patient) 95%
    c) Nurses responsive to my pain (in-patient) 95%
    d) Received adequate information before leaving (in-patient) 95%
    e) Staff was concerned about my comfort (out-patient) 95%
    f) Told me what to expect during experience (out-patient) 96%
2) About 2 monthly complaints due to unsatisfactory patient education by nurses in all nursing units at the hospital based on the "Patient Satisfaction Survey" space for comments.

Work level:

3) Nurses are currently able to perform patient education at the appropriate time. Nurses have the skills to apply the following four steps in performing patient education with 100% of the patients according to the following data:
    - Assess: Based on Nurse Questionnaire, 100% of respondents reported that they could assess the patient's need for patient education.
    - Plan: Based on Nurse Questionnaire, 100% of respondents reported they could plan a good education program for patients.
    - Implement: Based on the Nurse Manager Focus Group , nurse managers felt that nurses do an adequate job implementing patient education. Based on Nurse Questionnaire, 75% of respondents reported they could educate a patient every time they are with them.
    - Evaluate: Based on 10 patient education charts, when the patient education was documented within the chart, only 1 out of 15 (6.7%) instances of patient education was not documented as having been evaluated.
4) Documentation: Nurses do not always document every instance of patient education. They only document what they feel is important. They may document on one patient education sheet instead of on all the required patient education sheets. Based on the nurse questionnaire, only 12.5% reported that they document patient education every time they educate a patient. Based on the Patient/Caregiver Educational Needs Assessment sheet, under the nurses' notes sections of a patient's chart, 4 out of 10 (40%) needs assessments were documented.
5) Based on the nurse questionnaire, nurses ranked patient education as important as their daily tasks, 12.5%of nurses ranked patient education more important than their daily tasks, and 6% of the nurses ranked patient education as less important than their daily tasks. Based on the Nurse Manager Focus Group, nurse managers regard patient education as important as daily nursing tasks.
6) Patients are actually 94.7% educated concerning what patient education is and what their rights as a patient are regarding educational needs according to their responses to the Patient Satisfaction Survey. The three questions that relate from the survey are:
    a. I am able to care for myself (in-patient) 93%
    b. Received adequate information before leaving (in-patient) 95%
    c. Told me what to expect during experience (out-patient) 96%

Gap and Cost of Gap

Organization level:

1) Five percent discrepancy between the desired and actual rate of patient satisfaction based on patient satisfaction survey.
2) In reality, there are 2 monthly complaints, which is 2 over the desired 0 monthly complaints desired.

Work level:

3) Desired and actual performances are almost the same concerning dissemination of patient education.

4) Only 12.5% of nurses document patient education in every instances they educate patients. There is an 87.5% discrepancy between the desired and actual performances of patient education documentation. There is also a 55% discrepancy between the desired and actual workforce performance concerning the documentation of the needs assessment of patients documentation.
5) Actual performance meets the desired performance requirements for the nurses to regard patient education as at least as important as their daily nursing tasks.
6) Five percent discrepancy between the desired and actual levels of patients being educated concerning patient education and their rights as a patient.

Cost of Gap

The immediate cost of the performance gap is intangible and cannot be quantitatively measured in a short term. However, these intangible costs impact the overall reputation of the hospital at a macro-level.

First, the cost can be measured in a tri-annual report of patient education provided by JCAHO, poor documentation of patient education can lead to the hospital losing accreditation. A hospital without accreditation cannot operate at a level necessary to maintain a high reputation, or worst-case, remain in service at all.

Another situation that may arise and be a large cost to the hospital is if a patient was educated in a certain domain without documentation of the event. If a patient then suffered from mixing medications or another tragic event, and family members said that a staff member of the hospital never educated the patient regarding proper procedure of self-medication, no documentation exists to prove that the patient was educated in the aforementioned domain. From a legal standpoint, the education never took place, and the hospital's accreditation is at risk again.

Cause Analysis

Models

In determining the causes of the gaps, our team implemented Gilbert's Behavior Engineering Model. This model allows for identification of gaps and causes in both the environmental domain and the individual domain. Our PSA team felt that this model allowed for every possible cause present in the existing gap in patient education documentation at the hospital. Gilbert's model is very comprehensive, including six cells where three cells refer to the environment and three cells refer to the individual. One cell for each domain concerns *stimulus* events of the behavior, *response* events of the behavior, and *consequences* of the behavior, thus making up the 2*3 matrix. This addresses the antecedents to a behavior, the behavior itself, and the consequences of the behavior at both an individual and environmental level. Therefore, our PSA team felt this model was sufficient to identify causes of the identified gaps in workforce performance. It is assumed that the cause of the gap lies somewhere in one of the cells. It is possible that there are several causes in several of the cells. The prioritization process and results of this process will be dealt with later in this paper.

In applying the Behavior Engineering Model, Gilbert's statements in each cell were analyzed in accordance with the data that the PSA team collected through various methods as will be described in the next section of the paper. In applying the environmental cells of stimulus, response, and consequence, many factors that, if deficient, could be a cause of a gap, were discussed.

Our PSA team felt confident that these cause categories covered every possible aspect of human performance deficiency and would be a good stepping-stone in determining the causes of the performance gap at the hospital and identifying possible solutions.

Collected Data for the Cause Categories

Several data collection measures were implemented to gather information to address issues as described in PSA team project proposal. This information proved to be extremely valuable in completing information in the Behavior Engineering Model by Gilbert. The data collection methods that were implemented to obtain this information included nurse interviews, a nurse manager focus group, and several interviews with the Patient Education Manager. In addition, patient charts were analyzed, along with a patient satisfaction survey. The data collected from each of these methods helped in addressing the information in several cause categories. Although time constraints and patient privacy and rights prohibited all the data collection methods originally desired, the PSA team was satisfied with the amount and type of data that was able to be collected.

Environment information: A description of what is expected of performance is not explicitly stated, but it is known by all nurses. Through nurse interviews it became apparent that nurses were aware that they were supposed to document patient education.

It was also determined that there was no feedback given to the nurses about their performance or lack thereof.

Environment resources: There was a split in the advocacy of the newest grid implemented for use to document patient education. Some nurses thought the new grid for patient education documentation was easier than the old way, yet some felt it was too confusing.

Through nurse interviews, it also seemed that there was not enough time to document every instance of patient education.

Sufficient access to leaders was available; nurse managers and care coordinators (who passed the competencies on patient education) were available to all nurses as well as fresh sheets for patient education documentation.

The factor of sufficient personnel was also split. Some nurses felt more personnel would not help increase documentation while other nurses felt it would.

From our PSA team's point of view, work processes were not ideal to support patient education documentation.

Environmental Incentives: There were no financial incentives contingent upon performance of patient education documentation.

Minimal non-monetary incentives are provided but not contingent on documentation of patient education.

Career-development opportunities are not relevant to documentation.

There are no clear consequences of poor performance. This results in nurses not having to spend time documenting patient education and have more time to do other things.

Individual Knowledge: Nurses could document patient education if there lives depended on it. Lack of skills or knowledge is not an issue.

Training is provided although perhaps not sufficient or necessary. Most nurses feel that patient education is done to an adequate level. However, those same nurses feel it is not documented. The documentation of the patient education is not a complicated task that might require training.

Individual Capacity: Again, documentation is a simple task that is not being performed. There is no deficiency in the capacity of individuals to document patient education, therefore selection of employees based on this ability is not of concern. However, visual aids to augment performance may help remind nurses to perform.

Individual motives: It is unclear if nurses are willing to document patient education for incentives. Special motives may be required to encourage nurses to document the education they deliver.

Causes:

From the information above we can derive several causes of the performance gap in documentation of patient education. Four of Gilbert's cells have prompted us to recognize causes that will be addressed.

Stimulus: Environment Information does not provide feedback to the nurses on the adequacy of their performance.

Response: Environment Resources are not adequate. There are different forms in several units and not many are being used. Some nurses use the old sheet, some nurses complain about the new sheet, but most do not use any of the provided resources for documentation. Some nurses also stated that there was not enough time to document patient education. The work process for documenting patient education also prohibited documentation. One nurse suggested that if the chart was available to fill out in the patient's room directly after educating him/her, it would get done.

Consequences: *Environment Incentives* are not provided to the nurses for documenting patient education. Also, consequences for not documenting patient education are not present in the current system.

Consequences: Individual Motives to document patient education are not present.

These causes have been identified with the help of Gilbert's Behavior Engineering Model. In the next section of the paper, the causes will be prioritized in advancing toward identifying solutions to close the performance gap.

Cause Prioritization Process

Several causes of the performance gap were identified in the previous section. Prioritization of these causes is an important step and included the PSA team as well as the Patient Education Manager. The members of the PSA team and the manager rated the causes in terms of magnitude, value, and urgency, as defined by ePlan, (this is the EPSS to aid in the PSA and can be viewed at www.lpg.fsu.edu/eplan)

Magnitude was defined as an estimate of the level of the organization affected by this results gap.

Value was defined as an estimate of the cost of this gap to the organization in terms of direct and indirect losses.

Urgency was defined as an estimate of the acceptability of a delay in dealing with this result gap.

The results of team member's ratings along with the Patient Education Manager were averaged and placed in rank order.

Identified Individual Causes and Supporting Data

Based on the procedures mentioned above the causes are listed in rank order as follows:

1. Lack of feedback from managers about adequacy of performance
2. Lack of time to document the patient education chart
3. Using various forms to document patient education
4. No clear consequences of poor performance
5. No intrinsic motivation to document
6. No clear guidelines to document the patient education chart
7. No adequate incentives contingent upon good performance

After reviewing all the data that was collected, the PSA team noticed that there is no adequate feedback from managers about documentation on the patient education chart. In addition, the PSA team found the second gap from the nurse questionnaire in which 16 nurses provided answers. From the second nurse questionnaire, the PSA team discovered that each unit or each nurse uses different forms of a patient education chart because they may be used to the older form or some of them are not satisfied the most recent form. In addition, nurses do not feel that they have enough time to document everything they educate patients on. Nurses and nurse managers said there are no adequate consequences for not documenting patient education. Through the second nurse questionnaire, the PSA team learned that there are no specific guidelines or criteria for which to document the patient education chart. In the next section the PSA team will propose solutions to reduce the gap by introducing solutions in alignment with cause categories as defined by Gilbert's Behavior Engineering Model.

Recommended Solutions and Interventions

Process Used to Identify Solution

Our PSA team used the 'Possible Solution Generating Worksheet' from the ePlan tool to help with the process of identifying appropriate solutions to the causes of the previously stated gap between the desired and actual work performance. Our PSA team first had to identify the cause type of each of the four causes from the 'Matching Causes with Possible Solutions' worksheet also from the ePlan website. After identifying the cause types, the PSA team then identified the solution categories for each of the four causes using the 'Solutions Matched with Causes Worksheet'.

Matched with Causes Worksheet' to determine which ones were appropriate for each of the causes as potential solutions. Based on the 'Solutions Matched with Causes Worksheet' the PSA team determined that 17 (about 4 per cause) solutions were potentially appropriate. It was decided that the list of 17 possible solutions had to be narrowed down to a manageable number before they could be prioritized. Collaboratively, the PSA team decided to combine some of the solutions to narrow the list from 17 to 9 (about 2 per cause) possible solutions. Below is a chart listing each of the four causes and the original 17 possible solutions.

Cause	Cause Type	Solution Type	Possible Solutions
1) No Feedback from manager about adequacy and performance	Environment Information	Feedback System	Set performance expectations
			Set performance standards
			Production wall charts
			Performance appraisal
2) Lack of time to document	Environment Resources	Organization Redesign	Process redesign
			Improve data sharing
			Change reporting relationship
			Define or change job/goal responsibility
			Provide new or improved tools or technologies or work spaces
3) Using various forms to document	Environment Resources	Job Aids	Checklists
			Work Samples
		Organization Redesign	Participatory decision making
			Process redesign
			Provide new or improved tools or technologies or work spaces
4) No clear consequences of not performing	Environment Incentives	Recognition Systems	Non-monetary incentive
			Merit system
			Improved appraisal and recognition program

Solutions Prioritization Process

The PSA team then used the 'Solution Selection Rating Form' from the ePlan prototype to prioritize each of the 9 possible solutions that had been chosen. 'The Solution Selection Rating Form' had four major criteria categories: Opportunity, Capability, Collaboration, and Motivation. Opportunity was defined as the organization-level support and commitment. Capability was defined as the collective knowledge and skills of an individual, department, or organization. Collaboration was defined as the level of user involvement in adoption, adaptation, and implementation process. Lastly, motivation was defined as the perception or attitude potential adopters and stakeholders have of the attributes of the solution. A rating of 1 (disagree) to 5 (agree) was assigned to each of the questions under the major criteria heading by each of the PSA team members. Once each of the PSA team members had rated all 9 of the possible solutions based on the criteria from the 'Solution Selection Rating Form' the ratings were summed. The top five total team scores became the five possible solutions to close the performance gap.

Client and PSA team rankings of the five possible solutions to the performance gap, ranked in order:

1. Guidelines/ work samples. (189)
2. Performance and improvement appraisal
   Merit/demerit system---non-monetary incentives---recognition. (179)
3. Production wall charts. (173)
4. New tools/forms. (169)
5. Set performance expectations/standards, and define/change job/goal. (160)

Our PSA team made the suggestion to implement and evaluate the recommended solutions, but the hospital stakeholders will make all final decisions regarding any solutions that will be implemented.

Recommended Solutions and Their Relationship to the Causes and the Performance Gap

The final recommended solutions, derived from the process stated above, were made in rank order as shown above.

As previously identified, the gap that our PSA team was trying to close was the 87.5% discrepancy between the desired and actual behavior of nurses documenting patient education. Our PSA team identified the causes using Gilbert's Behavior Engineering Model. These causes are directly related to the gap, and the solutions directly related to the causes.

The recommended solution of *Guidelines and work samples* was derived from the EPSS ePlan prototype. This solution was derived from the environmental resource cell of Gilbert's Behavior

Engineering Model (BEM). The cause included lack of proper tools, which was noted from the data. Several forms were in use and some nurses did not like any of the forms that are used to document patient education. Our PSA team felt that a work sample or guidelines would remind the nurses of the correct way to use that sheet. A sample of the proper way to use the patient education documentation grid could be posted on the wall, on the front of patient's chart, or in front of 'Nurses' notes', a section within the patients' chart where the grid is found. Not only would this serve to remind the nurses the correct way to use the form, but remind them that this is an important task to complete every time it is performed.

The recommended solution of *Performance appraisal/improve appraisal leading to a merit/demerit system with non-monetary incentives, including recognition* is suggested as a result of the cause in the environmental information cell in Gilbert's BEM. Data showed that a lack of feedback from nurse managers existed with concern to the documentation behaviors of nurses. ePlan suggested a feedback system including a performance appraisal system. The PSA team felt that a performance appraisal system coupled with a merit/demerit system with non-monetary incentives being rewarded to exemplary performers by gaining several merits and recognition could increase the motivation of nurses to document the education of patients.

Ganong and Ganong, in their 1984 book, Performance Appraisal for Productivity: The Nurse Manager's Handbook mention several reasons and criteria for developing and implementing an effective performance appraisal system. One of these reasons is that it helps satisfy basic human needs (Maslow, 1970). A good performance appraisal system will focus on the strengths of individuals, allowing them to develop their shortcomings into strengths through self-actualization and self-appraisal. Performance appraisal systems, if developed and implemented correctly can foster a productive and efficient work environment.

Production charts posted on the wall near a high traffic area would be implemented to increase documentation of patient education to show nurses how they are doing. Hopefully nurses will take pride in seeing their performance line increase and approach 100%. This solution was derived from ePlan as well, concerning the same cause as described above. Nurses not receiving feedback from managers may be influenced to increase their documentation habits if their performance can be seen on a daily or weekly basis. Self-monitoring of this behavior is an option considering the amount of time it would take a nurse manager to synthesize all the charts and give graphic feedback to all nurses on their performance.

Graphic feedback coupled with task clarification has proved to be an effective intervention to increase desired behaviors in a variety of settings Crowell, Anderson, Abel, and Sergio (1988) found that task clarification and feedback increased the occurrence of appropriate customer service behaviors in bank tellers.

New tools (forms) was rated highly by all members of PSA team as using various forms was seen as a cause of the gap of poor documentation of patient education. This again was in the environment resource cell of the BEM by Gilbert (1978). However, a new form was recently implemented without data being taken to determine if an increase in documentation occurred; regardless a gap still exists. Currently, the hospital is slowly implementing a new computerized documentation process for all information and tasks, including medication and patient education. A nurse manager was also working on developing a new form to document patient education although time constraints restricted our PSA team from meeting with her. Our PSA team believes this could be a good solution if nurses had ownership in the development of the form; it may be used more often. However, our PSA team felt we could not offer assistance in developing a new form and it would occur without our input.

Our PSA team noted setting *performance expectations/standards and defining/changing job goals* was a possible solution. It, again, was derived using ePlan from the environment information cell of Gilbert's BEM. This solution also looked to solve the problem of absent feedback from managers to the nurses about their performance. Perhaps nurses would increase the documentation of patient education if the expectations were explicitly stated and standards were set.

Many of these solutions can be used together to reduce or close the gap in the documentation of patient education. Our PSA team's goals are to suggest solutions that will do just that. And we believe that the recommended solutions, when implemented correctly, would reduce or close the stated gap.

Implementation

In the context that our PSA was conducted, it was known that solutions would be recommended at the end of the semester. Implementation, if done, would be conducted by the stakeholder/client, in this case the Patient Education Manager would make arrangements if she saw fit. Our services were offered to answer any additional questions or meetings. At a time during the next semester we contacted our client to get an update on implementation, if any had occurred. The Patient Education Manager had implemented many of our recommendations, or some version of them. Some implementations were still being developed but the most visible of the solutions was a guideline, which clarified what was expected and how to fill in the patient education grid. In addition a 'Patient Education Newsletter' was developed for monthly distribution in paper version and on hospital intranet. The newsletter contained information on the need for patient education and one issue even had a graph with performance on pain management education efforts.

Again, our PSA team was not involved in the implementation effort so an evaluation of techniques used was null. However, we saw merit in the efforts and wanted to evaluate what had been put in place as addressed in the next section.

Measure the Impact of the Recommended Solutions

Evaluation is essential to determine the success of the implementation of performance solutions. Evaluation provides information on the effectiveness of the solution proposed by our PSA team on several levels. In our proposal, we provided a model for evaluation on our proposed solutions based on Kirkpatrick's four levels of the evaluation. Kirkpatrick's model has been widely used for training evaluation since 1970s. Our PSA team adopted this model according to our particular performance improvement solutions for the Learning Systems Department at the hospital. Kirkpatrick's original evaluation model is defined as:

Level 1: Measure the participant's reaction material presented.
Level 2: Measure if the participant learned the material presented in training (or other performance improvement program).
Level 3: Measure the participants' change in behavior as a result of training (or other program).
Level 4: Measure the impact of training (or other program) on the effectiveness of the organization.

Our PSA Team adopted this model for our evaluation efforts. Evaluation at all four levels was not possible with our limited time, resources, and access to necessary hospital information. In lieu of our proposed evaluation plan at the beginning of the analysis process, we concentrated our efforts on level 1 by conducting surveys on reactions to implemented interventions. The survey/interview was similar to the one implemented in the gap identification and cause analysis phase. In addition we asked questions in reference to the availability of the implemented interventions and the reaction nurses had to these. Level 3, change in behavior was conducted by looking at patient charts to determine if documentation had increased. In both evaluation efforts we visited the same units as for gap identification and cause analysis. Below are results of our evaluation efforts.

Results

Method: A Chi-Square test was used to determine if significant results were in place. We randomly selected four units of the hospital during analysis and returned to these units for evaluation. For each unit, 4 patient charts were randomly selected for examination. Each patient chart record has at least one documented event, at most ten daily records were observed. The same data collection procedure was repeated after the implementation of our PSA recommendations. We compared the patient education daily records before and after the implementation.

	Before Implementation of Our PSA Recommendations	After Implementations of Our PSA Recommendations
Sum of the Days Patient Education Recorded	10	45
Sum of Days Patient Education Not Recorded	31	42
Sum of Days	41	87

Fisher Exact Test Show the Probability $P<.001$. We then concluded that our intervention statistically improved nurses performance on patient chart documentation.

Discussion and Conclusion

We were pleased with our efforts in the analysis portion of the project, as was the patient education manager and our professor. We would have liked to have more time to help in the implementation of our recommendations such that we could have done it more systematically and regulated it. Regardless, we were satisfied with the results that occurred due to our efforts and those of the Patient Education Manager. We would have liked to have a larger sample size in our analysis and evaluation efforts, but were contented nonetheless.

Our main goals, in conducting the analysis, included learning the performance analysis process and to help the patient education efforts at the hospital. Our PSA team, our professor and the patient education manager felt that we had done both adequately.

In conclusion, our analysis determined that the main gap was in the documentation of patient education. Resulting from that, our PSA team recommended that Learning Systems consider the following interventions to close the performance gap on the nurse documentation process.

- Give nurses Guidelines/ work samples.
- Give nurse performance and improvement appraisal Merit/demerit system---non-monetary incentives---recognition.
- Show production wall charts.
- Use new tools (forms in particular).
- Set performance expectations/standards, and define/change job/goal.

Our PSA team believes that the success of these solutions in closing the gap at the hospital depends on the effectiveness of the implementation and the monitoring of these solutions. Our PSA team also believes that stakeholders play a crucial role in the implementation of the recommended solutions.

Regardless of our beliefs in implementation strategies, the solutions that were implemented yielded significant results which will serve the hospital well in future endeavors and future JCAHO visits (and we learned a lot thanks to this class and practical experience). We would like to thank our professor for expanding and contributing to our knowledge of the subject of the performance systems analysis process and the patient education manager for the opportunity to gain real world experience on our way to being professionals in the field.

References

Crowell, C. R., Anderson, D. C., Abel, D., & Sergio, J. (1988). Task clarification, performance feedback, and social praise: Procedures for improving the customer service of bank tellers. Journal of Applied Behavior Analysis, 21, 65-71.

Kirkpatrick, D.L. (1994). Evaluating Training Programs: The Four Levels. San Francisco, CA: Berrett-Koehler.

Maslow, A. (1970). Motivation and Personality, 2nd ed., Harper & Row.

Ganong, J. M. and W. L. Ganong (1984). Performance appraisal for productivity : the nurse manager's handbook. Rockville, Md., Aspen Systems Corp.

Gilbert, T.F. (1978). Human competence:Engineering worthy performance. New York: McGraw-Hill

Palm Education Pioneers: Examining the Potential of the Handheld Computer

Karen Swan
Mark van 't Hooft
Sebastián Díaz
Kent State University

Introduction

Current research shows that computer use and student learning gains are "closely associated with having computers accessible in teachers' own classrooms" (Becker, Ravitz, and Wong 1999; see also Marx, et al., 2000; Norris & Soloway, 2001; Soloway, Norris, Blumenfeld, Fishman, Krajcik, & Marx, 2001), and that a 1:1 student to computer ratio is needed to make computing in schools truly personal. For many school districts, especially the larger and poorer ones, attaining this ratio is a financial impossibility (see Norris & Soloway, 2001). Handheld computers, which cost a fraction of the price of desktop and laptop computers, can fill this gap. This technology may provide schools a more realistic alternative for integrating technology into the classroom and meeting the challenges of improving student achievement.

In May 2001, SRI International awarded nine Palm Education Pioneer (PEP) research hub grants to provide Palm computers for teachers and their students and evaluate uses of Palm handheld computers in K-12 classrooms (http://www.palmgrants.sri.com/background.html). The Research Center for Educational Technology (RCET) at Kent State University (KSU) was awarded one of these grants. As a research hub, RCET monitored the use of about 280 Palm IIIc devices in the classrooms of eleven teachers in ten local schools. The teachers who took part taught grades 1-12 in a variety of areas including computer science, math, language arts, integrated curriculum, and special education. RCET PEP staff developed and facilitated training sessions, provided technical and curricular support, and gathered and analyzed data to help evaluate the project. The PEP project was designed to address the following questions:

- How can handheld computers in the classroom improve teaching and learning?
- What educational activities do handhelds make possible?
- Just as important, what is missing in this new technology?
- How can handhelds be adapted to harness their full potential in the classroom?

To assist each PEP Research Hub in addressing these questions, SRI provided researchers the templates for surveys to be administered to both teachers and students utilizing the handheld computers in their classrooms. Researchers at RCET made the necessary modifications to these surveys and administered them to PEP students and teachers at the end of the school year (May 2002). The results of these surveys are addressed in this report.

Methodology

Subjects

Eleven teachers (grades 1-12) in a variety of subject areas including computer science, math, language arts, integrated curriculum, and special education participated in this project. They represented ten schools in six school districts in Northeast Ohio. Over 200 students participated in the project. As expected, not all students completed surveys due to attrition and lack of parental consent. In addition, not all teachers returned the end-of-project survey.

Survey Instruments

SRI developed and provided the surveys to be administered to both teachers and students at the end of the school year. For the purposes of this report, slight modifications were made to these instruments. The following data were collected for students in addition to the items in the SRI student survey: student's grade level, student's gender, and whether survey was self-administered by student or by proxy.

Data Collection

Most student surveys were administered during a two-week span at the end of May 2002. The investigators for this project traveled to the various schools to administer the surveys to either groups of students or individual students in cases where they needed one-on-one assistance. For some students who were absent during these administrations, teachers administered the surveys afterward and then mailed the completed forms to RCET. Teacher surveys were collected throughout May and June 2002.

Data Analysis

Quantitative data from the student surveys were analyzed using SPSS version 10.0. These analyses include a computation of frequencies for each item, as well as an exploratory analysis comparing male and female responses to each item. The qualitative items of the student survey were analyzed using a general thematic content analysis.

The teacher survey, although more extensive than the student survey, was completed by nine out of eleven teachers. Given this very small sample size, the quantitative results of that survey are not presented in this report.

Limitations of Study

This study has several limitations that should be considered when making decisions based on the findings. First and foremost, this study represents a preliminary investigation of how handheld computers can be utilized in the classroom to improve teaching and learning. Although they are not conclusive, the suggestions gained from teacher and student surveys, as well as teachers' suggestions in their journals and one-on-one feedback provide valuable insights that may be incorporated into future hardware and software design related to handheld computers. Other limitations include:

- The teachers were selected because they had some technology experience; they were not a random sample. Findings, therefore, may not be true of all teachers.
- The manner in which surveys were administered varied by classroom. Some classrooms were very quiet, for example, others were the opposite. In addition, special accommodations were made for some students to help them complete the surveys. This may have affected consistency of student answers.
- The student survey was administered to children in grades 1-12. However, according to the Flesch-Kincaid grade level measure available through Microsoft Word, the student survey in this study is written at approximately a 5th grade equivalent (4.7) reading level. The validity of this survey instrument may therefore be compromised by its reading level, especially for grade levels 1-5, which represent the majority of students (76%).

Results for Student Surveys

The Student Questionnaire was administered to approximately 230 students. Of these completed surveys, 217 are used in the data analysis reported here. Student surveys for which researchers were unable to obtain appropriate consent were not included in the study. One hundred and fourteen (53%) of students completing the survey were male, and 103 (47%) were female. One hundred and eighty-one (83%) students completed their own surveys, while 16 (17%) received one-one-one assistance from a researcher in order to complete the survey.

Student respondents ranged from 1st through 12th grade. At the elementary level, respondents included: 1 first grader, 25 second graders, 44 third graders, 17 fourth graders, and 78 fifth graders. In the middle school grade levels, respondents numbered 19 in both the seventh and eighth grades. Nineteen were in grades 9-12. The majority of children (84%) reported having never used a Palm before.

Student Use of Handheld Computers

All but two students reported using their Palms in their classrooms. About 34% reported using them elsewhere in the school building, including 12% who reported using them in a computer lab or the library. Five percent of the responding students said they used them at an outside study site, such as on a field trip. Twenty-three percent of the students reported using their Palms in outside locations other than study sites and 28% said they used them at inside locations other than school or home, such as on the bus or in a restaurant. Seventy-three percent of the students reported using their Palms at home. When asked how frequently they used their Palms at school, respondents were pretty evenly split between using them one to times a week (29%), three to four times a week (37%), and using them almost every day (32%).

Table 1 gives the percentages of students who reported using their handheld computers for a variety of activities in school and at home. Seventy-five percent of the respondents reported taking their Palms home. The percentages reported for home use are given for this population. The most frequent uses both in school and at home were playing games, using the calculator, drawing, and scheduling. In school, beaming was another frequent activity. Most of the least frequent activities reported involved peripheral devices – taking pictures with the camera, using the science probes and downloading games – but graphing software was only used one third of the time in school and very little out of school.

	at school	at home
taking notes	55	31
scheduling assignments	74	64
writing down things to remember	68	68
beaming	83	39
drawing pictures	72	74
using calculator	90	84
using graphing software	34	17
using science probes	25	13
taking pictures with the camera	25	12
downloading games	31	31
playing games	89	98

Table 1: *Percentage of Students Reporting Using Palm for Various Activities in School and at Home*

Students were also asked which functions they had used on their Palms. Most reported using the Memo Pad (92%), Calculator (92%), Infrared Beaming (88%), and games (90%). About two-thirds of the students reported using the Date Book (66%) and a little more than half reported using the To-Do List (61%), Desktop/Hot Sync function (58%), the Alarm/Timer (55%) and the Address Book (54%). Less than a third of all students reported downloading applications onto their Palm computers (29%). Relatively few students (9%) reported using Email.

Student Attitudes Toward Handheld Computers

The Student Survey asked students to respond "very," "some," or "not at all" to six questions regarding the use of handheld computers for learning. Table 2 summarizes those responses. It shows that most students found the Palms easy to use, fun, and a useful tool for learning. Almost all indicated that they would like to use them in the future. On the other hand, the majority of respondents thought desktop computers were better tools for learning.

	very	some	not at all
easy to use for schoolwork	58	39	3
makes learning more fun	64	32	4
helps you learn	58	35	6
helps you stay organized	41	46	13
want to use again	80	17	3
desktop computer better	23	41	36

Table 2: *Percentages of Students Indicating Various Degrees of Agreement with Questions about Learning with Palms*

Students were also asked to respond to each of 10 statements by indicating to what extent they agreed or disagreed with the statement (Strongly Disagree, Disagree, Agree, Strongly Agree). When asked to reflect on the statement, "I liked using a Palm computer at school," 94% of students agreed or strongly agreed. The majority of students (90%) either agreed or strongly agreed with the statement, "Using a Palm made learning more fun." Relatively few students (6%) agreed or strongly agreed with the statement, "Using a Palm computer is a big waste of time." Approximately 78% of students agreed or strongly agreed with the statement, "Palm computers are cool because they are kind of like an electronic toy." About half of the students (55%) agreed or strongly agreed that, "Using a Palm computer makes me a better student." Approximately 70% of students agreed or strongly agreed that, "Every student should have a Palm computer." About half the students (45%) agreed or strongly agreed that "A desktop computer is better than a Palm computer." Approximately 54% of students agreed or strongly agreed that "For taking notes it is easier to use paper and pencil than to use a Palm computer." The majority of students (85%) agreed or strongly agreed with the statement, "I wish my parents would buy me a handheld computer." The majority of students (82%) also agreed or strongly agreed with the statement, "I felt special because I got to use a Palm computer at school."

Open-Ended Questions

In addition to the quantitative items, the student survey also contained 5 open-ended questions. These were designed to augment the information sought through the quantitative items. Student responses to each of these items are summarized below.

The two more common themes that emerged from students' responses to ***"How does the handheld computer help you in school or in learning?"*** were organization and fun. Many students commented that the Palm handheld computers helped them to be more organized with aspects of their schoolwork. One student stated that "The handheld computer helps me store some information instead of giving it to someone else and then losing it." Another student said that "It helped me in school because it got me more organized than before when I didn't have one." A third student mentioned that "It helped me become more organized because I had a place to write things down."

Students also commented often on how the use of Palm computers made learning more fun: "The Palm helped me in school because it makes learning much more fun and made me pay a lot more attention." Some students also mentioned that the Palm computers helped them with their math and spelling. Examples of Palm use for math are, "The calculator helps a lot because we don't have a lot of calculators in the classroom"; "The calculator helped with my multiplication problems"; "It helps me graph stuff"; and "It helps me with my math homework." Examples of Palm use for writing include, "You can see your mistakes a lot easier then you could see it on paper. It helped a lot by teaching us a new way to learn"; and "For me, the Palms made it easier to find my mistakes when I was writing."

Others students mentioned that a particular feature or a particular software program was helpful. Examples of this include Palm features, such as the "Clock helps me sit down more than I was"; and using the Memopad to "Take notes and search for keywords in them." In addition, students mentioned the usefulness of a variety of software programs including Tessellation ("It helps in math "), Fling It ("Flinging a reading site and reading that and finding vocabulary"), Planetarium ("It helped me learn about space because my teacher put a space program on our Palms"), a downloadable periodic table for chemistry, and a dictionary.

Not surprisingly when asked ***"What is the best thing about using a Palm computer?"*** students most often identified the fun they experienced using them ("It was fun to do. It made math more fun. It made me learn something in a fun way"). This often translated into a higher level of motivation for learning. One student said the Palms "made it fun to learn about some things", and another stated "I was more anxious to learn and write things, and remember things, because I got to do it in the Palm." In addition, students often referred to how much they enjoyed the games available to them. Other advantages of Palms that were mentioned by students included their size ("If I had my own I would carry it everywhere"; "it is tiny with a lot of capabilities"; "You can put it anywhere (storage)"; and finally "The best thing about a Palm is that it is handheld").

Students were also asked to ***"Describe what things are not easy with the handheld computer, or that don't work well with the handheld computer."*** They identified a variety of challenges they faced with their Palm computers. One of the more common themes that emerged from students' comments was the difficulty in using the Palm computers to take notes. Students mentioned that "Writing was difficult for me," or "It takes long to write something." In some cases, students identified the use of the on-screen keyboard as the difficulty encountered because "If you press a letter it would change to a different letter." Another student mentioned the same problem, "Writing without the keyboard [was difficult] because sometimes the Palm misreads the letter you want it to be and makes it a different one." In many cases, this was resolved by recalibrating the screen. In addition, a few students mentioned the keyboard was "too small" because they were "used to a regular keyboard." In other cases, the difficulty of learning and using Graffiti was mentioned. Students thought that "Writing in graffiti takes too long and is complicated"; "Some letters don't look the same [like in the cursive alphabet]"; and "The letters were hard to remember."

Students also mentioned that drawing with the Palms was very difficult. In some cases they referred to a specific piece of software used for drawing such as TealPaint. Other pieces of software students mentioned having difficulty using included TeamLab, AvantGo, and the built-in alarm. Finally students mentioned problems with beaming and hotsyncing. A significant number of students mentioned having difficulty with their Palms not being digitized or calibrated.

Students also identified as a problem the need to keep Palm handhelds charged. On its face, this particular complaint may seem frivolous, but we need to remember that school use of handhelds is different from individual use. First, adult users usually each have their own cradle where they charge their Palm, while in a typical classroom, twenty or more students share four or five cradles. Secondly, typical adult use of a handheld computer outside the cradle may not be as energy intensive as students completing an hour-long exercise in class.

When asked ***"What would you like to change about the Palm handheld computer?"*** many wrote that they would like to see Palms offer more interesting options, especially in the areas of games, and "sound or music." Also, students mentioned that the Graffiti should be "more like real letters." Many students also identified the need for a bigger on-screen keyboard or a larger screen in general ("The thing that would make it better would be the screen getting clearer and bigger"). Other students suggested adding hardware features like voice recognition, a built-in digital camera, 3-D imaging, and faster processing speed. Finally, students expressed the desire for Internet access and email ("If you could email people on their computer").

Comparisons by Gender

Exploratory analyses of student survey responses revealed statistically significant differences in student responses to three items on the survey when comparing students by gender. The first of these items asked students if they had ever used a Palm computer before starting to use one in school during this project. A higher proportion of males (~ 20%) responded "yes" as compared to females (~8%). Another item that yielded a significant difference between the sexes asked students whether they had used a Palm computer in a computer lab or library. Again, a higher proportion of males (~ 18%) used their Palms in a computer lab or library as compared to females (~ 4%). Finally, gender differences were found in the degree to which students thought using the Palm handheld computers made learning more fun. Overall, females felt more strongly that the handhelds made learning more fun. For females, approximately 2% responded "not at all," 26% responded "some," and 72% responded "very." For males, 5% responded "not at all," 36% responded "some," and 57% responded "very."

Results of Teacher Surveys

Although the teacher survey contained many quantitative items, only nine teacher surveys were collected. This section of the report, therefore, will focus on the qualitative comments teachers made in their surveys. These results are presented in narrative fashion, and address the major themes found in comments in teacher surveys.

Projects

Due to the variety of grade levels and subject areas in the PEP hub at KSU, teachers and students were involved in a range of different projects using handhelds. Examples of units taught include a simulated stock market crash (5th grade integrated curriculum), polygons and nature (5th grade integrated curriculum), calculating the cost of a garden using spreadsheets (1st-3rd grade Special Education), family history, (8th grade Language Arts), collaborative story writing using beaming (2nd-5th grade Special Education; 2nd grade), collecting and analyzing data related to plant growth (7th grade math), and astronomy (5th grade science). In addition, several teachers mentioned that student collaboration increased due to the beaming capabilities of handhelds. One teacher mentioned that "students were able to work on various portions of a joint project and download peer work", while another one stated that his students used beaming during the writing editing. A second grade teacher mentioned that her class used Palms to write and share weekly math problems. By the end of the year they "had 552 math problems to look at and share with the entire class." Finally, one 8th grade teacher mentioned that students who used Palms to edit writing according to proficiency standards "scored significantly better than others."

In addition, students were involved in projects that reached far beyond their classrooms. Several classes took Palms on fieldtrips to places such as the Great Lakes Science Center in Cleveland, or to study landforms using the Kodak PalmPix cameras and Go 'n Tell software. A 4th/5th grade class met with members of the local Rotary club to demonstrate Palm use in their classroom. During this meeting, "students showed how they use date book and address book as personal organizers, graffiti for note taking, Fling-It for storing web pages, HandySheets for off site data collection, and AvantGo to keep up to date on local and national news." Some of the high school students in the project even used the Palms on the job.

Teacher Organization

Handhelds helped teachers to be more efficient and organized, and having a variety of information stored in one device was helpful: "I was just amazed at how dependent I became – I use it for everything I do!" Many teachers used Palms to keep track of student information such as grades or parent information. In addition, it was "easy to share information between home computers and those at school. Storing info on the Palm made it much easier to transfer work done at home to the classroom." Distributing information to students was simple as well, and would take no more time than it would take to pass out papers. Teachers would beam information to a couple of students, who in turn would beam it to others, until everybody in the room had the item beamed to them. Finally, one teacher used the Palms extensively to communicate with parents.

Nevertheless, some teachers raised logistical issues related to using a classroom set of Palms. One teacher in particular noted that she wanted to "be able to more closely monitor what the students are doing on the Palms" and in general have more control over how Palms were used by students. In addition, she mentioned the importance of administrative and technical support. In one district, the installation of Deep Freeze, a program that prevents users from modifying the hard drive on a computer, made hotsyncing, downloading programs, and getting Internet sites on the Palm "very cumbersome or impossible." Three of the teachers in the PEP hub were in this district.

Learning Curve

Teachers did not think they (or their students) had real difficulties learning how to use handhelds, with the exception of graffiti. Time was an issue, and most teachers stated that they needed more time to "truly explore all possibilities," and share with other teachers using Palms. In addition, several teachers mentioned the need for more training.

Pragmatic Suggestions

Most teacher suggestions are related to handheld hardware. For one, the displays caused several problems. Teachers in lower grade levels often said that the displays were too small and that this caused problems with tapping or writing. In addition, several teachers had problems with students putting too much pressure on the displays, either because of how they stored their Palms, or how hard they pressed on the screen. The majority of handhelds we sent back to Palm for replacement experienced this problem. Also, the fact that the displays on the Palms used were made out of glass caused irreversible damage that could have been prevented had they been made out of plastic. One teacher asked if something could be put on a handheld that "would alert students if they are applying too much pressure to the screen."

Other suggestions deal with the logistics of using a classroom set of Palms. One teacher mentioned that it should be easier to charge or sync multiple Palms simultaneously, and that it would be nice to be able to beam to multiple Palms at once (Bluetooth capabilities allow for that now). In addition, this teacher wanted to have more control over what students could and could not do with Palms, especially when they were shared between classes: "multiple students need to be able to store files on the Palm without other students being able to view, change, or delete them."

Access to Hardware and Software

Availability of hardware was somewhat of an issue, especially when it came to text input. Many teachers described how the availability of keyboards would have alleviated student frustration with text input either by using graffiti or the on-screen keyboard. A few teachers mentioned that other peripherals such as cameras or probes would have been nice, but money was usually not available for these.

Availability of software was by far the greater concern. Almost all teachers mentioned the need for more educational software in a variety of areas: "I felt the Palms were not used enough because I had difficulty finding good, useful applications;" and "The current software is business related. Far more educational software apps will need to be developed." Teacher wishes included a spreadsheet with Excel capabilities for math, a graphing calculator that does tables, multimedia presentation programs, grading/record keeping programs, and software that would allow them to lock certain features with passwords. While most of these programs are available in one form or another, they are often too expensive, especially if multiple licenses have to be bought.

Conclusions & Recommendations

Improving Teaching and Learning/Possibilities

In sum, how can handheld computers in the classroom improve teaching and learning? Our evaluation discovered that handhelds have a great potential in educational settings. Teachers and students agreed that “accessibility of a computer for each student is the greatest benefit. Students are able to collect, store, and organize data. They can research, calculate, write, and share information ... The Palms enhance student collaboration and encourage students to use higher level thinking skills.” Handhelds also enable students to take the initiative and explore, allowing for “a more authentic and deeper immersion in technology, not as a separate curriculum, but as an integrated part of [the] whole curriculum,” Moreover, handhelds encourage “the use of technology in everyday activities and enable students to understand the computer as a tool.”

What Is Still Missing?

Just as important, what is missing in this new technology? As mentioned, teachers and students ran into a variety of hurdles across the way. In addition to issues mentioned earlier (such as hardware durability, screen size, text input, syncing problems, logistic issues, easier Internet access, and a lack of affordable software), there is a need for more standardized connections for cradles and peripherals, and more compatibility with Apple computers. The latter is an important point because a large number of school districts use Apple computers (3/10 in our project).

Modifying Palms for Classroom Use

Originally, handheld computers were designed with the individual adult consumer in mind. Providing tools such as a calendar and address book, they were seen as eventually replacing more traditional planners. Today, handhelds are everywhere, including in classrooms. As such, hardware and software designers need to focus their attention more on developing hardware and software aimed at K-12 teachers and students. The main issues PEP teachers and students ran into during their participation were hardware related. First, due to the fact that we used an older model Palm (IIIc) which uses a glass screen, about 5 Palms were lost due to cracked screens. Next, a substantial number of Palms (about 15%) had to be replaced because they would no longer recognize the stylus. In general, this was due to too much pressure on the screen for too long a time. Some of this was resolved by the purchase of Palm cases for. Finally, in a few cases the limited amount of memory became a problem.

In addition to issues related to hardware, teachers voiced concerns about the availability of software. The IIIc’s used in this project came with the Palm OS and the basic functions. However, once teachers realized the potential uses of their handhelds in the classroom, they wanted more software. Unfortunately, they often did not have the time or money to acquire what they needed, and this is where PEP hub staff stepped in to help.

In sum, while widespread use of handhelds in K-12 education can potentially have a substantial impact on teaching and learning, hardware and software development need to be geared more toward education. The ideal handheld package for the K-12 classroom would include a handheld with a plastic display, rechargeable battery, and expandable memory. In addition, this handheld would come with Bluetooth capabilities, a preloaded bundle of educational software geared toward specific grade levels (e.g. elementary school), and have wireless Internet access. Finally, this handheld would come with a keyboard and a sturdy case.

Implementation Strategies for the Classroom

Teachers mentioned a variety of issues when discussing the use of a classroom set of handhelds. The issue that was by far the most pressing was that of management of a classroom set of Palms. Teachers had to figure out ways to effectively deal with loading new software, control over what went on the Palms, keeping the batteries charged, and general Palm maintenance. In several classrooms, students tended to make the screens as bright as possible, with the result that they ran out of battery power in a couple of days. Keeping the batteries charged has also become an issue during the summer time, as teachers who do not want to reinstall software at the beginning of the new school year will need to keep them charged over summer break. Recalibration of the screens was also a big issue.

Another major issue related to the use of classroom sets of Palms had to do with breakage, repair, and warranties. In educational settings, Palms are used more often and suffer more abuse at the hands of students than they would in the hands of the average adult consumer. The initial warranties are only good for one year, and when schools make substantial investments in technology, they should not have to buy additional warranties after one year, or replace broken equipment at full cost. Many school districts simply cannot afford either.

A related issue to breakage is that of loss or theft. Since the Palms used in the PEP hub are the property of individual schools, each school had to come up with its own policy. Many teachers had parents sign a form stating that students were going to use handhelds and that they would be responsible for them. In one instance, parents said they would only sign the forms if their children used the handhelds only in the classroom. Consequently, this group of children never took the Palms home. In one fifth grade classroom, two students lost their Palms. One was found, the other was replaced by the parents. In another fifth grade classroom, two Palms were stolen. They were returned, but only after the teacher offered a $10 reward with no questions asked. All in all, however,given the extensive use of the handhelds in our PEP hub, relatively few of them were broken, lost, or stolen. It appears that regular maintenance and parent involvement were essential in keeping the Palms in good running order.

Future Research

Sex Differences Issues

An exploratory analysis of the student survey data yielded a significant statistical difference when comparing boys' and girls' responses on several items. These findings suggest that gender differences may exist in access to and perceptions of handheld technology. More research needs to be conducted to further these differences.

Special Education

In this study, at least two Special Education teachers used handheld computers with their students. The use of handheld computers in this environment raises interesting questions about how they impact this type of student. For instance, while some teachers expressed concerns about the use of handheld computers and the impact on students' penmanship skills, one Special Education teacher commented on how the Palm helped her to conduct spelling exercises that would normally be a rather daunting task given student learning disabilities related to reading and writing. More research needs to be conducted in this area

Assessment

Handheld computers have a great potential for addressing the challenges of student assessment in the classroom. Particularly in those cases where assessment need not be of the high stakes variety, students can use Palms as a tool for self-assessment. There are several potential advantages to doing this. First, the administration and scoring of such assessment becomes easier for the teacher, who can beam and/or download student work to their own handheld or to a central database. Second, the Palm computer may incorporate software features that would help make the assessment process more fun and authentic, thus encouraging students to engage in self-assessment. Third, software programs exist that provide both student and teacher a graphical and easy-to-read display of student progress in a particular subject area over time. To facilitate testing on handheld computers, however, much research needs to be conducted to determine what challenges and concerns may arise in classroom settings. Such research might address the effects of self-assessment on student achievement. It might also examine more pragmatic issues, like difficulties that may arise when administering assessments using handheld computers.

Studies of specific subject areas and/or grade levels

Now that we are beginning to see the overall picture of what it means for teachers and students to have access to a ubiquitous tool such as a handheld, research needs to be done in specific subject areas and at different grade levels. For some areas and at certain grade/age levels handheld use may be more suitable or appropriate than for others. Very little research has been done in this area, and what is available is mostly anecdotal.

References

Bannasch, S. (1999, Fall). The electronic curator: Using a handheld computer at the Exploratorium. *Concord Consortium Newsletter* (http://www.concord.org/library/1999fall/electronic-curator.html).

Becker, H., Ravitz, J.L., & Wong, Y. (1999). Teacher and teacher-directed student use of computers and software. Report #3, Teaching, learning, and computing: 1998 national survey. Irvine, CA: Center for Research on Information Technology and Organizations, University of California, Irvine.

Marx, R. W., Blumenfeld, P., Krajick, J., Fishman, B., Soloway, E., Geier, R., & Tal, T. (2000). Inquiry based science in the middle grades: Assessment of student learning in the context of systemic reform. Ann Arbor: University of Michigan Center for Learning Technologies in Urban Schools. Unpublished manuscript.

Norris, C., & Soloway, E. (2001, June). Towards realizing the potential of palm-sized computers in K-12. iMP Magazine. http://www.cisp.org/imp/june_2001/06_01soloway.htm.

Pea, R. D., & Gomez, L. (1992). Distributed multimedia learning environments: Why and how? *Interactive Learning Environments, 2* (2), 73-109.

Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology, 1* (1), 145-168.

Soloway, E., Norris, C., Blumenfeld, P., Fishman, B., Krajick, J., & Marx, R. (2001, June). Devices are ready at-hand. Communications of the ACM. http://www.handheld.hice-dev.org/readyAtHand.htm.

Tinker, R. (1997, July 7). *The whole world in their hands*. Concord Consortium (http://www.concord.org/library/pdf/future.pdf).

What Types of Technology Do Middle and High School Mathematics Teachers Really Use?

Berhane Teclehaimanot
William Weber
University of Toledo

Abstract

The purpose of this study is to examine the extent to which middle and high school mathematics teachers use of technology in the classroom and to what degree their students are using technology. Today many states have established standards for the use of technology in the classroom including the state of Ohio. To answer many of the questions surrounding technology in schools, a study was conducted with randomly selected middle and high school mathematics teachers in Ohio.

Introduction

The use of technology in public education continues to grow as we enter into the 21st century. As a result, many schools have been wired for internet access in the last three years alone (DeMedio & Teclehaimanot, 2001). The integration of technology into the curriculum has become a catchphrase in K-12 education. In one major national report on school technology readiness, it was pointed out that all students must graduate with the technology skills needed in today's world and tomorrow's workplace and all educators must be equipped to use technology as a tool to achieve higher academic standards (CEO Forum, 1999). According to the findings from Apple's Classroom of Tomorrow, effective technology use does enhance student learning (ACOT, 1995); but unfortunately half of the teachers were not utilizing the technology available in their schools and some were not using computers at all. In addition, the National Center for Educational Statistics (2000) reports that nearly all public school teachers reported having computers available for teachers somewhere in their schools. Although most teachers recognize the integration of technology into the curriculum requires new approaches to the teaching and learning process, not enough consideration is being paid to ensuring that teachers who are teaching with the new innovative technologies have been adequately trained (Rowe, 1999). In order for teachers to integrate technology across the curriculum, there must be institutional support that rewards teachers with incentives. In addition, there must be a support infrastructure available to help them to troubleshoot and solve technology related problems without any delay and interruption of classroom activities.

Teachers' adoption of technology into the curriculum is key in transforming teaching and learning in the classroom. To highlight this importance, The National Council of Teachers of Mathematics established a Technology Principle in its Principles and Standards for School Mathematics (NCTM, 2000). NCTM states that "Technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances students' learning" (p. 6). The NCTM envisions that every student has access to technology that is used to help facilitate the learning of mathematics. The use of technology can help students develop a deeper understanding of mathematics by providing visual representations, helping organize and analyze data, and allowing exploration of important mathematical ideas.

According to Maney (1999), it is not the technology that makes the difference, but rather how teachers adapt and apply technology. There are many factors that influence teacher's utilization of technology in the classroom. Rogers (1995) stated that a positive attitude toward any innovation increases the likelihood of adoption of the innovation. Teachers' beliefs and attitudes toward technology, fear factors or complexity issues, lack of time and lack of support, limed access, lack of faculty development training and lack of organizational support have been identified as major barriers. There are other factors that block the inclusion of technology in the classroom such as lack of computers in the classroom, access to the Internet, outdated software and hardware, and lack of release time for teachers to learn how to integrate technology into the curriculum. While numerous studies have examined the use of technology in the K-12 classrooms, few have taken the focus one step further to determine the extent, to which technology is used and required in the K-12 education.

Accordingly, most educators today believe that public schools must ensure the effective use of technology in the classroom in order to prepare our children for 21st century learning environments. For that reason, the U.S. Department of Education's "Preparing Tomorrow's Teachers to Use Technology" (PT3) and the International Society for Technology in Education (ISTE) implementation grant projects are playing an essential role in disseminating information to ensure that our future K-12 teachers are technology-wise and can use the multimedia facilities available in today's schools. There are many questions regarding the uses of technology in schools. Are teachers adequately prepared to use technology in their teaching? Are schools providing appropriate types of technology for teachers and students? What types of technology do teachers really use in their mathematics classrooms? There also are major questions dealing with the role of technology in schools. For example, what are the goals of using technology in elementary, middle and high schools? How can technology improve student learning in mathematics? Middle and High schools, in particular, can play an important role in helping students become skilled in the use of technology. Not only can technology help middle and high graders learn mathematics more effectively, its use can also enable students to improve their technology skills. To answer many of the questions surrounding technology in schools, nine hundred (900) middle and high school mathematics teachers in the State of Ohio were sent surveys as part of this study. These teachers were identified as middle and high school mathematics teachers (grades 4 – 12) through data obtained from the Ohio Department of Education database system. A total of 232 surveys were returned yielding a response rate of 26%. This study used a self-reported survey instrument to explore the types of technology used by middle and high school mathematics teachers, the technology use by their students, and the facilities and equipment provided by their schools. The questionnaire consisted of two parts. In Part One, the teachers marked how often they used particular types of technology and also the types of technology used by students. Part Two of the survey asked about the availability of different types of technology in their schools and if their school provided professional development training workshops to integrate technology into the mathematics curriculum. In addition, teachers were ask to indicate what kind of support would encourage them to utilize technology in their mathematics classroom.

Demographic data on the survey included the number of years teaching experience, grade level taught, age, gender, the size of the school the teacher taught in, and if the school was urban, rural, or suburban. Teachers were assured that confidentially would be maintained. The results were analyzed using percentages of technology use in the classroom. The majority of those responding were female and most had one to ten years of experience. Moreover, most of the respondents were ninth grade mathematics teachers and most were from suburban and rural schools system.

Results

Frequencies were calculated on all data collected. All results were reported as percentages. The results in Table 1, show that the most commonly used type of computer platforms in the mathematics classroom were; PC or PC compatible computers 68%, Apple Macintosh 32%; and the most used computer platform in the schools' computer laboratory were 77% PC or PC compatible computers and 20% Apple Macintosh computers. Furthermore, 58% of the respondents indicated they have only one computer in the mathematics classroom for instruction. The most frequent forms of technology used were: Internet access for teaching 89%, e-mail access to discuss teaching with other educators 91%, computer access in the teaches workroom 46% and 97% of the mathematics classrooms were wired to the Internet.

Table 1: *Types of Computer Platforms Used by Mathematics Teachers and Access to the Internet*

	% Used	Yes	No
PC (Windows) or PC compatible computers in the classroom	68%		
Apple Macintosh computers in the classroom	32%		
PC (Windows) or PC compatible computers in the lab	77%		
Apple Macintosh computers in the lab	20%		
1 computers in the mathematics classroom	58%		
2 computers in the mathematics classroom	13%		
3 computers in the mathematics classroom	7%		
No computers in the mathematics classroom	6%		
Computers in the teacher workroom		46%	54%
Mathematics classroom wired to the Internet		97%	3%
Internet access in the mathematics classroom		89%	11%
Email access in the mathematics classroom		91%	9%

Results reported in Table 2 show the use of technology in the mathematics instructions by teachers and students: almost every day 14%, once or twice a week 16%, once or twice a month 30% and never or hardly ever use technology for instruction 40%. In addition, the data shows the use of technology in planning instruction almost every day 20%, once or twice a week 28%, once or twice a month 2%, and never or hardly ever 31%.

Table 2: *Technology Used by Mathematics Teachers and Students*

	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Never or hardly Ever
Use of technology in mathematics instruction	14%	16%	30%	40%
Use of technology in planning instruction	20%	28%	21%	31%

The results of this study indicate several important findings concerning technology in middle and high schools. Table 3 shows the most often used types of technology were: word processing for teaching 68%, Microsoft PowerPoint for teaching and presentation 64%, spreadsheets for assisting teaching 57%, database to assist teaching 32%, computer graphics for teaching 28%, and the of HyperStudio for teaching 27%.

Table 3: *Technology Used by Mathematics Teachers*

	Yes	No
Use of word processing for teaching (MS Word)	68%	32%
Use of word processing for teaching (MS Works)	34%	66%
Use of word processing for teaching (Apple Works)	32%	68%
Use of word processing for teaching (WordPerfect))	18%	82%
Use of word processing for teaching (Other))	0%	100%
Use of spreadsheets for assisting teaching (MS Excel)	57%	43%
Use of spreadsheets for assisting teaching (Apple Works)	57%	43%
Use of database to assist teaching (MS Access)	32%	68%
Use of database to assist teaching (MS Works)	17%	63%
Use of database to assist teaching (Apple Works)	28%	72%
Use of filemaker pro to assist teaching	10%	90%
Use of computer graphics for teaching (Apple Works)	28%	72%
Use of computer graphics for teaching (MS Woks)	21%	79%
Use of computer graphics for teaching (PageMaker)	20%	80%
Use of computer graphics for teaching (Adobe Illustrator)	12%	88%
Use of computer graphics for teaching (Aldus Super Paint)	2%	98%
Use of Microsoft PowerPoint for teaching	64%	35%
Use of HyperStudio for teaching	27%	73%

Results reported in Table 4 show the types of technology required or encouraged by teachers for their students: Require students to use word processing for assignment 62%, encourage power point use by students 57%, encourage students use spreadsheets 47%, encourage students use of database 25%, and encourage students to use computer graphics 23%.

Table 4: *Technology Encouraged or Required for Students*

	Yes	No
Require students to use word processing	62%	38%
Encourage power point use by students	57%	43%
Encourage students to use computer graphics	23%	77%
Student use of spreadsheets	47%	53%
Student use of database	25%	75%

Findings in Table 5 indicate that schools are providing many types of technology for their teaching. CD-ROM drives are provided for a majority of teachers 77% and access printers is provided for a majority of teachers 70%. Digital camera equipment is available in nearly 64%, scanners 59%, along with digital camcorder 30%, while VCR's are available for teacher 32%, and Video (LCD) projectors are available in several schools 42%.

Table 5: Mathematics Teachers feel the following Equipment Should be Available for Teaching:

	Is Available	Should be Available
Scanner	59%	25%
Digital camcorder	30%	20%
Digital camera	64%	15%
CD-ROM drive	77%	6%
CD-ReWritables	15%	20%
DVD-RAM ReWritables	13%	16%
Printer	70%	8%
VCR-Videocassette recorder/player	32%	41.1%
Document camera (elmo)	14%	9.0%
Video projector (LCD)	42%	18%
New computers	34%	36%

Table 5: *Do Your School District Provide Professional Development Workshops to Implement Technology to Teach Mathematics?*

	Is Available	Should be Available
Scanner	59%	25%
Digital camcorder	30%	20%
Digital camera	64%	15%
CD-ROM drive	77%	6%
CD-ReWritables	15%	20%
DVD-RAM ReWritables	13%	16%
Printer	70%	8%
VCR-Videocassette recorder/player	31.7%	41.1%
Document camera (elmo)	14%	9.0%
Video projector (LCD)	42%	18%
New computers	34%	36%

Discussion

Evidence from this study indicates several important findings concerning technology in middle and high school mathematics classroom. The study supports the concept that given the opportunities for teachers to collaborate and to work as teams to develop technology integrated lesson plans, the availability of updated hardware, software and Internet accessibility in the classroom, continued administrative support, professional development to use new technologies and incentives, the majority of teachers will demonstrate an interest in learning and using technology in the mathematics classroom. In additions, this finding is important because it supports the views of a majority of educators and the public who want more technology in schools. It is interesting to note that there is a significant difference in response rate from teachers in suburban schools versus urban or inner-city schools (suburban response rate was twice that of urban).

The results of the analysis of technology usage levels indicated that mathematics teachers who had taught (1 to 5 years) were generally using newer technologies. Mathematics teachers with more teaching experience (26 to30 years) were generally the most uncomfortable with using newer technologies. Furthermore, significant differences in the amount of technology use were found between new mathematics teachers (1-5 years) and more experienced mathematics teachers (11-15 years). The reason why most of the new teachers use technology in their classroom could be due to state requirement on technology during their preservice teacher education programs.

Middle and high school teachers are using many types of technology in their classroom. Beside computers, teachers are using e-mail to discuss teaching with other educators and the Internet to locate resources and other materials to enrich their curriculum. The majority of mathematics teachers use word processing, spreadsheets, graphic calculators, power point, CD-ROM's, digital cameras, and other contemporary form of technology. It is interesting to note that majority of teachers are integrating technology into the curriculum to enhance the teaching and learning process of their students. As the result, middle and high school teachers are encouraging and requiring that their students use technology in their work. This result, too, is important because it reinforces the belief that students need technology to aid their learning and to develop technology skills needed to excel in the 21st century.

To most educators, these findings may be particularly surprising. The data from this study demonstrated that 97% of mathematics classrooms were wired to Internet with one to five computers in the classroom and other equipment important in motivating teachers to use technology in the classroom. It is discouraging to note that, almost half middle and high school mathematics teachers never or hardly use technology for instructional purposes in the classroom. In addition, even though opportunities for professional development training workshops for learning new technologies were provided; some teachers did not take the advantage to attend the workshops.

Although the study indicates that some middle and high school mathematics teachers are using technology in their teaching, it is imperative that the use of technology becomes more prominent in all mathematics classrooms. To help make this a reality, professional development for technology use is critical and must be specific to teaching mathematics. In addition, advanced technology mediated classrooms, administrative support, and the availability of technical support to teachers is critical to encourage teachers to utilize technology in their teaching of mathematics. It is also clear that all mathematics teachers need to continue and increase their use of technology in their classrooms. Students too, must continue to improve their skills in using technology by using it for their learning. Schools, also, must continue to provide different types of technology, resources, facilities and training for both teachers and students. As we enter the 21st century, it is hoped that all middle and high school teachers will make technology an essential part of their teaching and that they will strive to help their students become "technologically literate."

References

Apple Classroom of Tomorrow. (1996). Changing the conversion about teaching, learning & technology: A report on 10 years of ACOT research. Cupertino, CA: Apple.

CEO Forum. (1999). The CEO forum school technology and readiness report. Professional development: A link to better learning. Washington, DC: CEO Forum.

DeMedio, D. L., & Teclehaimanot. B. (2001). What types of technology do middle school teachers really use? Michigan Middle School Journal, 26 (1).

Education Week. (1998, Oct. 1). Technology Counts '98 [Special issue]. Education Week. 28 (5).

International Society for Technology in Education (ISTE). Eugene, OR. http://www.iste.org

Maney, J. K. (1999). The role of technology in the systemic reform of education: Reality, pitfalls, and potential. In G. Cizek (Ed.), The handbook of educational policy (pp. 387-415). San Diego, CA: the Academic Press.

National Council of Teachers of Mathematics. (2000). Principles and standards for school mathematics. Reston, VA: NCTM.

Rogers, E. M. (1995). Diffusions of innovations (4th ed.). New York, NY: The Free Press.

Rowe, D. (1999, Oct/Nov). Teaching with technology: Faculty development approach. Community College Journal.

U.S. Department of Education, National Center for Education Statistics. (2000). Teachers' tools for the 21st century: A report on teachers' use of technology. Washington, DC. U.S. Government Printing Office.

Motivating Learners in Educational Computer Games

Hakan Tuzun
Indiana University Bloomington

Abstract

The purpose of this study was to identify motivational elements for an online multiplayer educational computer game. Design ethnography with naturalistic interpretations has been used as a research method. After the data analysis by using the constant comparison method of grounded theory, thirteen categories emerged for the kids as the motivational elements to play this game: identity presentation, social interaction, playing, learning, ownership and control, fantasy, immersive context, curiosity, creativity, achievement, rewards, uniqueness, and context of support.

Introduction

Even the most elegantly designed educational software will fail if the learners are not motivated to learn. For this reason, designers of any educational software must try hard to establish a context that learners will find motivating. Motivation is related with learning because learning is an active process which requires conscious and deliberate effort (Bruner, 1960; Wlodkowski, 1986). Even the most skilled learners are unlikely to learn if they do not apply some effort. In recent years there has been an extensive effort in building learning environments that provide motivational elements. In particular, some educators are examining the potential of computers and even using principles underlying the design of video games to establish rich learning contexts (Barab, Thomas, Dodge, Carteaux, & Tuzun, in press; Malone, 1980; Lepper & Malone, 1987).

Theoretical Background

The framework for motivation in educational games has been largely provided by Lepper and Malone's (1987) work (Tzeng, 2001). They developed a taxonomy of intrinsic motivations for learning, based on four factors as motivating to the learner: challenge, curiosity, control, and fantasy.

While informative for the design of educational contexts, Lepper and Malone (1987) examined video games for entertainment. There has been relatively little research that examines the use of gaming context as providing an educational context for supporting learning. Additionally, as Weiner (1990) points out, motivation should not be limited with 'the self.' Since learning is a shared activity and since it does not take place in a vacuum, new motivation theories should incorporate new constructs like 'belongingness' and 'cooperative learning.' This study will research the motivations of users interacting with a learning and teaching project, Quest Atlantis (QA) that uses a 3D multi-user environment to immerse children in educational activities.

Study Background

Quest Atlantis Description

Quest Atlantis is an educational computer game that immerses children in a virtual environment for completing educational activities. A detailed description of this game was articulated elsewhere (Barab et al., in press). Below, a brief introduction is provided to familiarize the reader to the study background. The purpose of Quest Atlantis game is to save mythical Atlantis from an incoming disaster. According to the back story of the game, as the Questers complete the educational activities called 'Quests', they help with saving the Atlantis from this disaster. Quest Atlantis has foundations on three bases: education, entertainment, and social commitments.

However, instead of conceptualizing Quest Atlantis just as simply a computer software, it might better be described as a virtual environment designed to support an online community as well as multiple face-to-face communities. The Quest Atlantis storyline, its virtual worlds, and policies make up the Quest Atlantis meta-game, a term used in the commercial gaming sector. The Quest Atlantis meta-game contains the following key components:

- A mythological legend that provides a back story for Quest Atlantis activities
- A number of 3D worlds and villages through which Questers, mentors, and the Quest

Atlantis council members can interact with each other
- A Personal Digital Assistant (PDA, a kind of homepage) for each Quester, serving as a portfolio of their learning and participation
- An advancement system centered on pedagogically valid activities that encourage academic learning, entertainment, and social commitments
- Extrinsic rewards structure

Quest Atlantis combines play, role playing, adventure, and learning, allowing learners to immerse into virtual 3D worlds where they select or they are assigned developmentally-appropriate Quests, talk with other Questers and mentors, and build virtual persona (Turkle, 1995; Bers, 2001). Quest Atlantis is implemented in different contexts, including schools as part of the curriculum through QA unit plans, and after-school programs as a volunteer activity (i.e., Boys and Girls Clubs of America).

Quest Atlantis has many components that can be categorized under different major groups: for example, communication, collaboration, and ownership. Within the game the modes of communication are chatting in the 3D space, the internal e-mail system, telegramming, and other discourse within the physical space through various means (i.e., talking within the computer lab, or Questers' talking through phone). The modes of collaboration are co-questing, being part of a guild, requesting help from others, and helping others related with different QA tasks. The modes of ownership are having a personal PDA with various elements on it (emoticons, awards, etc.), X-points that Questers accrue after successfully completing Quests, having a unique avatar through customization, renting virtual land and building on it, artifacts created as the result of the Quests, and merchandise (QA trading cards, QA rulers, QA pencils, etc.) that can be purchased from Quest Atlantis trading post in exchange of the X-points.

Methodology

This study employs qualitative research paradigm. In this sense, the framework of this study can be characterized with multiple labels. It can be characterized as an ethnographic research since its purpose is to describe a group (Fetterman, 1998). For example, I spent considerable amount of time among the people at the selected research site. In addition to one year of frequent visits prior to this study, I spent two months at the site for data collection. Since "good ethnography requires both emic and etic perspectives" (Fetterman, 1998, p. 22) I tried to capture both the insider's and outsider's perspectives of reality. The study included common elements of ethnographic studies such as field work, participant observation, and interviews.

This study can also be characterized as a naturalistic research study (Lincoln & Guba, 1985), because the data collection took place in a natural setting. Also, there were no variables manipulated to confirm or disconfirm a priori hypothesis.

However, being one of the designers of this educational game complicates my role as a researcher in addition to those challenges traditionally associated with ethnography or naturalistic research (Clifford & Marcus, 1986; Fielding & Fielding, 1986; Silverman, 1993), because I have been more than a participant observer. The philosophy of the Quest Atlantis implementation calls for collaboratively developing a vision for each of the centers, while this vision is researched at the same time it is created. Barab, Thomas, Dodge, Newell, and Squire (2002) refer to this process as 'design ethnography,' which is a design work that involves sustained participation and interaction with a context and a designed product. It is design ethnography that drives the methodology of this study.

Design ethnography draws upon the work of a collection of methods, including ethnographic research and action research. Action research, also called as participatory research (Adelman, 1993), has emerged when positivist paradigms failed in studying human organizations (Susman & Evered, 1978). Susman and Evered (1978) argue that the cyclic process of action research eliminates the limitations of the positivism and deals with practical concerns of the people. Actions are planned in mutual agreement by the researcher and the researched. Action researcher acts as a catalyst, who helps to the researched by handling problems and offering interventions to those problems (Hart & Bond, 1995).

In this sense, the goal of design ethnography can be described as changing or empowering the culture under study (Barab et al., 2002). During the process individuals and local contexts transact in a co-evolving fashion. The researcher wears the hat of a change agent (Rogers, 1995) and his goal is to support a transformational process. While doing so his role is a mixture of "peripheral membership" (Adler & Adler, 1987, p. 36) and "active membership" (Adler & Adler, 1987, p. 50) (i.e., he is both outside the culture and within the culture).

Research Question

The purpose of this study is to identify motivational elements for an online multiplayer educational computer game. The above methodological efforts were guided by the following research question:

> What are the motivational elements of Quest Atlantis, whether intrinsic or extrinsic, in terms of student-defined motivation?

Since motivation is a hypothetical construct (Martin & Briggs, 1986; Good & Brophy, 1997) and differs among academicians we need to define the motivation from the perspective of this study. In this study, motivation is defined as individuals' showing their willingness to initiate and sustain participation in Quest Atlantis activities. Examples of QA activities are completing Quests (engaging curricular tasks), participating in synchronous and asynchronous discussions, joining guilds, and signing up for QA jobs.

Context Selection

I conducted a purposeful sampling for the selection of the context. The following were the two major criteria for identifying the context: (a) Quest Atlantis centers that enrolled in Quest Atlantis program in the same region that I lived in, and (b) Quest Atlantis centers that enrolled in the Quest Atlantis program for at least six months. As a result this study took place in an after-school program located in a Midwestern town. The details of how the Quest Atlantis design and research team collaborated with this center has been articulated elsewhere (Barab et al., 2002).

Participant Selection

The participants of this study were members of a Boys and Girls Club meeting the following criteria: (a) Quest Atlantis players, who have played the game at least five different sessions, and (b) Quest Atlantis players, who have spent at least three hours within the game. With these criteria I wanted to make sure that the Questers have accumulated the prerequisite skills necessary to play the game at a basic level. I conveniently interviewed a total of twenty participants. As it is typical in Midwestern communities, almost all participants were Caucasian Americans. Just one of the participants was African-American. Additionally there were five female and fifteen male participants.

Data Collection Methodology

I used primarily ethnographic methods including interviews, observations in the different areas of the Club with an emphasis in the computer lab, and document analysis. Later, a demographics questionnaire emerged to support the interviews. Table 1 summaries these data collection methodologies.

Table 1. *Data Collection Procedure Summary*

Methodology	**Sources**	**Procedure**
Interview	Participants	Recorded semi-structured interviews on a digital voice recorder, then transcribed
Demographics Questionnaire	Participants	Interviewees filled in after the interview is completed
Observation	Observed computer lab Observed other areas of the Club	Took notes on scratch paper, recorded with a digital voice recorder and digital camcorder at times
Document Analysis	Materials available at the Club such as annual meeting reports and member information Materials available electronically on the Quest Atlantis servers	Read all materials and document any descriptive statistics related to interviewees

Data Analysis Procedures

Qualitative data analysis is inductive rather than deductive. The researcher starts with the data, and then develops concepts and categories, instead of beginning with theory, predicting a pattern of results, and examining the data to test the deduction. Therefore, instead of starting with a hypothesis, the researcher generates the hypotheses from the data. (Fielding & Fielding, 1986).

I used constant comparison method of grounded theory for data analysis. Grounded theory is a systematic set of methods to collect, code, and analyze data. (Glaser, 1992). Specifically the grounded theory is

> … A general methodology of analysis linked with data collection that uses a systematically applied set of methods to generate an inductive theory about a substantive area (Glaser, 1992, p. 16).

In constant comparison method, the researcher asks the following question while he continually codes, compares, analyzes, and writes memos about the data while analyzing them: "What category or property of a category does this incident indicate?" (Glaser, 1992, p. 19). The categories inductively emerge out of the data rather than being decided prior to the data analysis (Patton, 1987). Possible data sources might include interviews, field observation records, documents, and video tapes (Strauss & Corbin, 1994).

Glaser and Strauss, the inventors of the constant comparison method, originally described four stages to analyze the data through constant comparison method. These stages were (Glaser & Strauss, 1967, p. 105): (1) comparing incidents applicable to each category, (2) integrating categories and their properties, (3) delimiting the theory, and (4) writing the theory. However, in their later work they reorganized these into three stages. These stages are (Glaser, 1992; Strauss & Corbin, 1998): (1) open coding, (2) axial coding, and (3) selective coding. In their later work Glaser and Strauss chose to develop their own versions of the grounded theory and constant comparison method.

The approach I followed using the constant comparison method for my data analysis is neither Glaser's nor Strauss and Corbin's in their pure format. It can be said that I followed an adapted version of the constant comparison method by adopting tactics from both versions. Since Weiner (1990) pointed out to the 'many uncharted areas to incorporate' into motivational theories, I perceived the inductive approach in data analysis as a suitable tool that will enable me to unearth these 'uncharted areas.' Aligned with this idea I hesitated to force data as much as possible, and let the data emerge. For this reason, it can be said that my adaptation of using grounded theory and constant comparative method is closer to that of Glaser's (1978, 1992) in philosophy. Below I describe open coding, axial coding, and selective coding stages of my data analysis.

Open Coding. At this stage, data are broken down into its parts. While doing so, incidents are examined closely, and they are compared for differences and similarities. In this sense data are conceptualized. The purpose of conceptualization is to reduce mountains of data into manageable chunks by abstracting that data. While doing so, the labels might be assigned by the researcher, or the labels might be taken from the words of the participants. This latter case is referred as "in vivo" codes (Glaser, 1978, p. 70).
The open coding of the interview documents were done by three researchers and continued intensively for a whole week period. After the open coding of the interview documents there were 202 codes. Two-weeks of open coding of the observation records added 32 new codes. At the end we obtained a total of 234 codes.

*Axial Coding***.** This is the stage where categories are systematically developed and related to each other along their properties and dimensions. Open and axial coding are not sequential stages. The researcher keeps on coding for properties and dimensions while he develops relationships between categories. When it seems like no new properties or dimensions of a category emerge during the coding, that category is considered saturated. (Strauss & Corbin, 1998).

In this stage we tried to systematically develop the categories based on the codes. Open coding and axial coding were not sequential stages. We moved back and forth between the two stages. Three researchers got together to discuss the codes. Since the open coding was done independent of the research question, the codes were characterizing the data well in general but not respect to the research question of the study. Researchers debated the codes grounded in the data by using their own characterization of the motivation based on salient themes and the research question, and by visiting the current literature on motivation theories. So this debate was a dialectic intersection of the codes grounded in the data, our intuitive responses to the research question, and the current theories of motivation.

At the end we obtained thirteen categories: identity presentation, social interaction, playing, learning, ownership and control, fantasy, immersive context, curiosity, creativity, achievement, rewards, uniqueness, and context of support. Eventually, these thirteen categories were used to answer the research question.

Selective Coding. In selective coding the theory is integrated and refined (Strauss & Corbin, 1998). To this end the emerging story is explained around a core category. All other categories are linked to this core category. The core category accounts for most of the variation in the problem.

In this stage the main purpose was to obtain a core category and link other categories to this core category. A secondary purpose was to link the thirteen categories with each other. These two acts together and the categories grounded in the data helped with asserting my arguments. In this section the data were interpreted in a way that has both 'experience-local meaning' (Geertz, 1973) at the same time having 'experience-distance significance' (Geertz, 1973) to others analyzing motivation in other contexts and conditions. This occurred through ensuring that local interpretations were informed by and respond to previous research and theory, with presentations of the assertions being contextualized in terms of the broader literature. Limited space prohibits providing the outcomes of selective coding in this manuscript.

Trustworthiness

Triangulation has been a common method to provide trustworthiness in qualitative research (Patton, 1980). The term comes from the application of trigonometry to navigation. Locating the precise point of a geographic location requires using two points. The intersection of these two points gives the precise geographic location for the navigators (Bogdan & Biklen, 1998). Denzin (1970) originally advocated the use of multiple methods and multiple sources of data to provide triangulation in a qualitative research study. He later suggested using multiple researchers and multiple theories to improve the triangulation (Denzin, 1978; Merriam, 1998). I followed the first three approaches of triangulation to increase the trustworthiness of this study. I applied multiple methods to collect data: semi-structured interviews, observations, questionnaires, and document analysis. I used multiple sources of data: children using QA, personal observations, membership information in club records, and QA participation information in QA electronic databases and logs. And finally three researchers participated in the constant comparative analysis of data. Since I wanted to discover the 'uncharted areas' in motivation, finding any of them would contradict with the other theories of motivation. For that reason, I did not use other theories for the triangulation of my findings.

Findings

After the data analysis thirteen categories emerged for the kids as the motivational elements to play QA. These categories are identity presentation, social interaction, playing, learning, ownership and control, fantasy, immersive context, curiosity, creativity, achievement, rewards, uniqueness, and context of support. Each of these categories also has sub-categories. Below these categories are explained briefly and quotations are provided to illuminate the categories.

Identity presentation.

Kids present their identity through their avatar choices, homepages, and by their usernames. These three elements let them express themselves and show themselves to others. In this way they feel that they are empowered within the game.

> *Researcher: If you compare [QA] to Cartoon Network, how is it different?*
> *Interviewee: Well, it's more digital than Cartoon Network. You can interact with others across the world on both of 'em BUT it's more like you can see other people. You can see what…the avatar is part of it because it shows you what they like, what they are like.*
>
> *Researcher: What do you think about your homepage?*
> *Interviewee: Um, it's cool, I guess 'cause people can learn about you.*
> *Researcher: What are the things that you like on your homepage?*
> *Interviewee: Um…how you can say what you like and what you can do.*

Social interaction.

Social interactions happening in the game was one of the biggest motivators for the players. These interactions happened both within the online space and within the physical space where they connected to the game. While interacting with others they interacted with various people through multiple communication modes. At times there was competition among the Questers but sharing conquered over competition. Although there were Questers who wanted to play the game individually, playing it as a group was more frequent. And security features within the game bettered the social interactions.

> *Researcher: So Quest Atlantis created an opportunity for you to meet [new people]?*
> *Interviewee: I think Quest Atlantis is trying to make you friends. I think that's why the people made Quest Atlantis, so if you were just lonely at the club, you could get friends on Quest Atlantis. That could help you have more fun at the club.*

Playing.

To most Questers QA is a game that they can play. They name QA experience as 'playing the QA.' Multiple factors contributed to the gameness of QA. Controlling an avatar was one of them. Another factor was the different worlds and villages that made up the game. They explained that looking for Quests in them was an exciting activity to do. Exploring these worlds and villages and finding out secret places in them were fun things to do for them. They used their points in the trading post to buy store items. The synergy of all these factors made QA a fun game. Most of the interviewees fell in love with QA space, its characters, and its story:

> *Researcher: I know you like Cartoon Network pretty much right? How is [QA] different from Cartoon Network?*
> *Interviewee: Quest Atlantis is a thing that is only one thing, and you are actually a real person and you can talk to people and you don't have to just play a game, or anything. It's a learning thing where you can learn and have fun too at the same time.*

> *Researcher: Why did you join Quest Atlantis?*
> *Interviewee: I just thought it would be fun to play a game that you have to do quest and get points and you can build houses and stuff like that.*

Learning.

For Questers, QA is not just a game that they can roam around. This game also has an educational value and they enjoy learning through QA. Most of them emphasized learning in QA was a fun activity. They indicated QA was a place where 'you can learn and have fun too at the same time.'

> *Researcher: How is Quest Atlantis different from other things you do on the computer?*
> *Interviewee: You have fun while you're learning.*

This combination of fun and learning were such tangled that at times they were not aware that they were learning. In this sense they experienced flow (Csikszentmihalyi, 1990) in the learning process:

> *Researcher: How is Quest Atlantis different than other things you do on the computer? For example, I see a lot of kids, including you, play Cartoon Network. How is it different from Cartoon Network or from other educational games?*
> *Interviewee: Well, Quest Atlantis doesn't have that many games than Cartoon Network or any of that. It's different than Cartoon Network and stuff like that because it's got education, and you learn, and sometimes you don't even know that you are learning.*

Ownership and control.

The design and development of QA is based on a design model called 'Participatory Design.' In this sense the opinions of the users of the game are constantly evaluated by QA design team and reflected into the game. As the result of this approach, half of the Questers felt as if they were the rulers of the game most of the time. In this sense they treated the game as 'their own game.' The kids were definitely aware that QA was created by some outside people:

Researcher: Can you tell me about your favorite worlds or villages?
Interviewee: I like…Ocean World.
Researcher: What is the reason for that?
Interviewee: Uh…I just really like the ocean and it's cool that Quest Atlantis has got it where you can go down in the ocean and talk to people. 'Cause no other games would like, consider the ocean. And you guys, like, make us swim down there and stuff.

However, this creation by the outside people was in the form of implementing it. Actually most of these implementation ideas came from the Questers themselves. For example, when one of the Questers got a pack of trading cards I asked if he liked them. He pointed specific clutters on one of the cards and told me 'I designed these.' He was referring to the previous site visits by the game designers, who collected ideas from the kids related with card design. Therefore, he was implying to me that he liked the cards since he had a saying on their design.

Fantasy.
The fantasy elements the Questers like about QA include the QA myth and the QA council. At least half of the interviewees mentioned these two elements specifically.

Researcher: Can you explain more about why you complete quests?
Interviewee: Two things: I want to help the Atlantians, plus points.

Researcher: So, you like questing, you like the council…and what else can you tell me?
Interviewee: Yeah. Um also, when I make a mistake, I always like that because I like to read the letters that the council sends me if I made a mistake on my quest. I like to read those a lot.

Immersive context.
Although most of the action takes place on a computer screen, the QA experience utilizes other support structures, which elevates the game play and makes this experience an immersive one. These support structures at the club included QA posters, QA activity chart, QA trading cards, and QA comic books. 3D part provided most of the immersivenes on the screen.

Researcher: How is it different from other things in the computer lab?
Interviewee: It's different 'cause it's more interactive. It's more digital. You can, like, walk around in a digital space.

Researcher: Why do you think [QA] is fun to play?
Interviewee: Because, like, you get to do stuff, something like that, you don't get to do other things in there. It is like you are inside the computer. Because like that.
Researcher: How is Quest Atlantis different than other games in the lab?
Interviewee: It is different, because you get like, it feels like you are inside of it…

Curiosity.
There were several aspects within the game that made the Questers curious. These aspects increased their interest towards the game, and they wanted to come back to the game more often. These aspects were the end of the game, Quest response statuses, and secret places.

Researcher: Can you tell me those reasons [that make you come back to QA]?
Interviewee: I like the building stuff and also…I like doing quests and I just…last week I didn't even get to come to the club the whole week, so I was dreaming about going back to the club and seeing if the council has, um, accepted my quests yet, and they did. I was so happy about that.

Researcher: What makes Healthy World special?

Interviewee: Healthy World…it's got a lot of secret places. That's what I like about a world.

Creativity.

Questers said that they like 'creating stuff.' They like to be creative and they like to convert their ideas into reality. They all indicated again and again that the way to do this in QA was through building activities. In the virtual space they can build virtual structures, furnish them with different objects like pictures, and therefore impact the QA space. For one of the Questers, building became such an obsession that he kept working non-stop for three days to build his first virtual building. One Quester, who was observed to be obsessed with building, explained that QA let her to be creative:

Researcher: Why do you like building so much?
Interviewee: Because it's fun to like, make houses, be creative and make up a whole bunch of ideas and look at other people's houses too.

Eight out of twenty Questers chose building as one of their three most favorite activities in QA. In addition one fourth of the interviewees indicated that building was one of their reasons to come back to play QA.

Achievement.

While the Questers participate in different QA activities including doing Quests, building activities, finding secret places, and similar activities most of them perceive these activities as a challenge. They try to overcome these challenges. When they succeed their accomplishment is recognized by the game by different modes. Eventually they get satisfaction out of this overall process.

Researcher: You said you complete the quests for points, right? Is there another reason for completing quests?
Interviewee: To, uh, get land and just to have fun.
Researcher: What kind of fun are you having, because you already told me that some of the quests are hard to do and take a lot of time to type… So, what kind of fun are you having?
Interviewee: It's fun, like, the challenge to get it done just so I can get the points…it's just a challenge to get 'em. That's really my reason.

Researcher: What is the most exciting thing you have done in Quest Atlantis?
Interviewee: Like, whenever I go the idea that I could make a party room, and everyone started coming to my house and looking at it. And it really made me feel good.

Rewards.

There were two kinds of rewards that they recognized: awards on their homepages and material items. Material items included trading cards and other items that they could buy with their points. All kids indicated that they liked having the ownership of these rewards.

Researcher: What do Quest Atlantis points mean to you?
Interviewee: The points mean that, uh, if you get enough points you can either get some cards, or if you even get enough and go to this one special place in the trading post, you can get Internet time, a pencil, just basically anything that gets listed there.

Uniqueness.

Most Questers play QA because it is 'a game that sticks out from all the others.' It is unique because it creates a unique opportunity to do different things. In this way, it is different from other educational games, other computer software, and other activities in Questers' daily lives.

Researcher: Why did you join Quest Atlantis?
Interviewee: I thought it'd be pretty cool, because most of the Internet games are not learning environments, you know shoot'em all, kill'em kind of games. So yeah, I

thought it would be really cool to join something this totally different from that.

Researcher: Is [QA] different from worksheets [at the school], for example?
Interviewee: Worksheets you have to read, and on quests you also have to read. And on the worksheet, you have to write something down. On the computer, you just type it and it's like words. You send it by puttin' it in the inbox. I mean, at school you put it in the inbox. Here you just send it to the council. They'll read it, reply and, uh, give you your points. Like, one out of ten is for a worksheet.
Researcher: If you had a choice at your school. Your teacher came over, let's suppose, and said, 'Okay guys here is Quest Atlantis. Within Quest Atlantis you will complete this quest. And here is a worksheet.' Which one would you choose?
Interviewee: I'd say Quest Atlantis.
Researcher: Quest Atlantis? If there is no difference between them, why Quest Atlantis?
Interviewee: Because you get to change into someone and get to go to other worlds, but you can't go here; you can, uh, see the sites that have been provided for you; you can lift things, as in Healthy World you can lift a bike; you can see a big basketball; you can see other people's houses, what they've built; you can see pictures; you can build a house.
Researcher: With the worksheets you couldn't have those?
Interviewee: It's not really that much fun with a worksheet. Unless you have to like, change the worksheet into an airplane, then a paper ball, and then throw it into the trashcan!

Context of support.

Since QA was implemented in multiple contexts, including after-school environments and schools, it is noteworthy to point to the contextual implementation differences. Three of the interviewees also played the game in their schools and they all pointed to the differences in these implementations, which made the game play experience different. For example, at some schools they were not able to choose their usernames, but a username was assigned by a teacher. Another contextual difference was the variation of the trading post items for different schools.

Discussion

This study has found thirteen categories as a reason for the kids to play an educational computer game: identity presentation, social interaction, playing, learning, ownership and control, fantasy, immersive context, curiosity, creativity, achievement, rewards, uniqueness, and context of support.

When these categories are linked to make a sound interpretation towards a theory of motivation, three important constructs emerge: who participates, what activity is done, and what outcome is obtained.

Identity presentation and social interaction categories belong to 'who participates' construct. In this sense an individual likes to be part of a culture or social structure, however the individual still would like to keep his identity through various means. Therefore, an individual can participate in activities alone or he can participate with others.

Playing and learning categories belong to 'what activity is done' construct. For any activity, there were varying opinions on the type of that activity. For example, some kids characterized completing Quests as learning while some other kids characterized the same activity as playing. After the analysis, it was clear that most of the kids did not mind learning through this educational game although it was a volunteer activity at the after-school context. The intermingling of playing and learning contributed to this positive attitude. Based on this, it can be suggested that play is an important element for individuals and it should be mixed with learning or working to motivate them towards sustaining and completing these activities.

Achievement and rewards categories belong to 'what outcome is obtained' construct. Most of the kids liked having the rewards after they overcame the challenges. However there was still a minority enjoying the satisfaction out of challenges and refusing any extrinsic rewards.

Integrating these categories into a framework of motivation was in progress as of writing this manuscript.

Conclusion

This manuscript is part of a dissertation study. At the time of writing this manuscript, the writing of the dissertation was in progress. Limited space prohibits providing all the details of the study. Interested reader can find more information on this manuscript by looking into the same title under Dissertation Abstracts International (DAI), or by searching author's name under databases that index major publications in Social Sciences (when they become available). The author can be contacted through htuzun@indiana.edu or hakantzn@yahoo.com.

Quest Atlantis project (http://questatlantis.org) was supported in part by a CAREER Grant from the National Science Foundation, REC-9980081 and by the National Science Foundation Grant #0092831.

References

Adelman, C. (1993). Kurt lewin and the origins of action research. Educational Action Research, 1(1), 7-24.

Adler, P. A., & Adler, P. (1987). Membership roles in field research. Newbury Park, CA: SAGE.

Barab, S. A., Thomas, M. K., Dodge, T., Carteaux, B., Tuzun, H., (in press). Making learning fun: Quest Atlantis, a game without guns. To appear in Educational Technology Research and Development.

Barab, S. A., Thomas, M. K., Dodge, T., Newell, M., Squire, K. (2002). Design ethnography: Building a collaborative agenda for change. Submitted to Anthropology & Education Quarterly for publication.

Bers, M. (2001). Identity construction environments: Developing personal and moral values through the design of a virtual city. The Journal of the Learning Sciences, 10(4), 365-415.

Bogdan, R. C., & Biklen, S. K. (1998). Qualitative research for education: An introduction to theory and methods (3rd Ed.) Needhan Heights, MA: Allyn & Bacon.

Bruner, J. (1960). The process of education. Cambridge, MA: Harvard University Press.

Clifford, J., & Marcus, G. (Eds.). (1986). Writing culture: The poetics and politics of ethnography. Berkeley, CA: University of California Press.

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: Harper and Row.

Denzin, N. K. (1970). The research act: A theoretical introduction to sociological methods. Chicago, IL: Aldine.

Denzin, N. K. (1978). The research act: A theoretical introduction to sociological methods (2nd ed.). New York, NY: McGraw-Hill.

Fetterman, D. M. (1998). Ethnography: Step by step (2nd ed.). Thousand Oaks, CA: SAGE.

Fielding N. G., & Fielding, J. L. (1986). Linking data. Beverly Hills, CA: SAGE.

Geertz, C. (1973). The interpretation of cultures. New York, NY: Basic Books.

Glaser, B. G. (1978): Theoretical sensitivity: Advances in the methodology of grounded theory. Mill Valley, CA: The Sociology Press.

Glaser, B. G. (1992). Emergence vs forcing: Basic of grounded theory analysis. Mill Valley, CA: Sociology Press.

Glaser, B., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago, IL: Aldine.

Good, T. L., & Brophy, J. E. (1997). Looking in classrooms (7th ed.). New York: Longman.

Hart, E., & Bond, M. (1995). Action research for health and social care: A guide to practice. Buckingham, Open University Press.

Lepper, M. R., & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), Aptitude, learning and instruction: Conative and affective process analyses (pp. 255-286). Lawrence Erlbaum Associates: Hillsdale, NJ.

Lincoln, Y. S., & Guba E. G. (1985). Naturalistic inquiry. Newbury Park, CA: SAGE.

Malone, T. W. (1980). What makes things fun to learn? A study of intrinsically motivating computer games. Dissertation Abstracts International, 41 (05), 1955B. (UMI No. 8024707)

Martin, B. L., & Briggs, L. J. (1986). The affective and cognitive domains: Integration for instruction and research. Englewood Cliffs, NJ: Educational Technology Publications.

Merriam, S. B. (1998). Qualitative research and case study applications in education: Revised and extended from case study research in education (2nd ed.). San Francisco, CA: Jossey-Bass Inc.

Patton. M. Q. (1980). Qualitative evaluation methods. Beverly Hills, CA: SAGE.

Patton, M. Q. (1987). How to use qualitative methods in evaluation. Newbury Park, CA: SAGE.

Rogers, E. M. (1995). Diffusion of innovations. New York: Free Press.

Silverman, D. (1993). Interpreting qualitative data: Methods for analyzing talk, text, and interaction. London, SAGE.

Strauss, A., & Corbin, J. (1994) Grounded theory methodology: An Overview. In N. K. Denzin, & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 273-285). Thousand Oaks, CA: SAGE.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). Thousand Oaks, CA: SAGE.

Susman, G. I., & Evered, R. D. (1978). An assessment of the scientific merits of action research. Administrative Science Quarterly, 23(4), 582-603.

Turkle, S. (1995). Life on the screen: Identity in the age of the internet. New York: Simon & Schuster.

Tzeng, S. C. (2001). Optimizing challenges and skills in the design of an educational computer game and exploring adolescents' gaming beliefs. Paper presented at the meeting of the Association for Educational Communications and Technology Conference, Atlanta, GA.

Weiner, B (1990). History of motivational research in motivation. Journal of Educational Psychology, 82, 616-622.

Wlodkowski, R. J. (1986). Motivation and teaching: A practical guide. Washington, D.C.: National Education Association.

The Effect of Case-Based Versus Systematic Problem Solving in a Computer-Mediated Collaborative Environment

Daniel Uribe
United States Air Force Academy

James D. Klein
Arizona State University

Abstract

The purpose of this study was to investigate the effect of two instructional methods (case-based vs. systematic approach) and collaborative group size (dyads vs. quads) on learner performance in solving ill-defined problems in a web-based environment. Working in teams of two or four, participants used a web-based program that taught them a specific problem-solving approach. One program focused on how to solve problems using a case-based approach and the other program taught a systematic, four step problem-solving process. Participants then applied the problem-solving approach to solve realistic problem scenarios. Results indicated that participants who worked in dyads performed significantly better than those who worked in teams of four. However, those who worked in teams of four spent significantly more time on the web-based program than participants working in dyads. Analysis of time spent solving problems revealed that participants who used the systematic approach spent more time solving the problems than participants in the case-based approach. Participants in all treatment groups had positive attitudes toward working with others, the web-based programs, and applicability of problem-solving skills to other settings. However, results indicated that participants rated collaborating on-line the lowest. Analysis of the on-line communications between team members indicated that 95% of the interactions that took place were related to the problem-solving task. Implications for the implementation of computer-mediated collaboration in distance learning environments are discussed.

Introduction

Problem solving is regarded as one of the most important activities in everyday life and a primary goal of the education process (Jonassen, 2000; Phye, 2001). How to design effective programs that help learners acquire the skills to solve problems has become a focus of the instructional technology field in recent years. However, some critics argue that current instructional design models are not well suited to teach the skills necessary to solve ill-defined problems (Van Merrienboer, Clark, Moore & de Crook, 2002). To fill this apparent gap, several instructional strategies for problem solving based on current cognitive and learning theories have been advanced. These strategies emphasize solving realistic problems in authentic contexts (Oliver & Harrington, 2000; Bastiaens & Martens, 2000).

Proponents of case-based learning describe it as a powerful instructional approach that is engaging and that leads to sustained and transferable learning of problem-solving skills (Mergendoller, Bellisimo, & Maxwell, 2000; Stepien & Gallagher, 1993). Case-based learning uses authentic, complex problems as the impetus for learning and fosters the acquisition of both disciplinary knowledge and problem-solving skills (Bligh, 1995; Edens, 2000; Flynn & Klein, 2001; Levin, 1995). Julian, Kinzie and Larsen (2000) propose that case-based learning is more effective than didactic teaching methods because it more accurately represents the complexity and ambiguity of real-life problems and it provides a means for allowing learners to develop the kind of problem-solving strategies that practicing professionals use.

Van Merrienboer et al. (2002) identify case-based learning in the 4 Components Instructional Design (4C/ID) model as a scaffolding technique that supports the learner at the beginning of complex tasks. According to Van Merrienboer et al., case studies are product-oriented support for complex learning that provide the learner with a given state, a desired goal state, and the solution to the problem. They theorize that this scaffolding will help the learner develop mental models from the given examples that will aid in the solution of the same type of complex problems.

Another technique that can be used to solve complex tasks is that of a systematic approach to problem solving. Van Merrienboer et al. (2002) describe the systematic approach as a strategy that uses heuristics and/or rules-of-thumb, which can be applied to solve a complex task. Gagné (1985) describes problem solving as an extension of both rule learning and schema learning; he argues that problem solving is not just the application of rules to achieve a goal or solve a problem, but it is also a process that yields new learning. A systematic approach to solving problems has been found to be effective in a variety of academic disciplines (Benjamin & Hamdy, 1993; Van Streun, 2000).

A web-based distance education environment may be well suited for the teaching and learning of complex skills such as problem solving. In a web-based environment, a realistic setting can be created where the learner has access to a vast amount of authentic information and can work collaboratively through real-time computer-mediated communications to solve ill-defined problems. According to Laffey, Tuper and Wedman (1998), computer-mediated learning on the World Wide Web is suitable for problem-based learning because it provides ample resources, allowing students to do their own planning and present new forms of knowledge, which expand the mechanisms for collaboration and communication. Miller and Miller (2000) suggest that the characteristics of the Internet environment, which provides a hyperlink structure with easy access to relevant information, realistic and enhanced media and synchronous communication capabilities, make it an effective learning environment for complex skills. Others also argue that computer-mediated collaboration and the web are excellent technologies for case studies and integrating higher-order learning (Jonassen, Prevish, Christy & Stavrulaki, 1999).

Research suggests that a collaborative learning environment can positively affect performance on problem solving tasks (Flynn & Klein, 2001; Johnson, Johnson, & Smith, 1991; Mergendoller et al., 1999). Collaborative learning is defined as "an activity that is undertaken by equal partners who work jointly on the same problem rather than on different components of the problem" (Brandon & Hollingshead, 1999). Smith and McGregor (1992) also defined collaborative learning as a “joint intellectual effort.” A meta-analysis of the use of collaborative learning in higher education courses indicated that collaborative learning promotes higher achievement, higher-level reasoning, more frequent generation of ideas and solutions, and greater transfer of learning than individual or competitive learning strategies (Johnson et al., 1991). Other researchers suggest that collaborative learning is effective in areas requiring higher order thinking skills such as problem solving (Damon & Phelps, 1989; Doran, 1994).

Research supports the hypothesis that a collaborative learning environment is well suited for solving problems (Spector & Davidsen, 2000). In several studies conducted to analyze the impact of a collaborative environment on problem solving, collaboration was found to improve performance on complex or higher-order thinking activities (Chang & Smith, 1991; Johnson & Chungh, 1999; Mergendoller et al., 1999). In these studies, learners appeared to benefit from the opportunity to discuss the problem, brainstorm potential solutions and arrive at a final solution. However, these studies have been conducted in face-to-face environments. Additional empirical research is necessary to indicate whether the positive effects of collaborative learning during problem solving tasks will also be obtained in a computer-mediated collaborative environment. With enrollments in courses delivered over the Internet in the United States already at well over 100,000 students (Simonson, Smaldino, Albright & Zvacek, 2000), it is important to confirm that the positive effects of collaborative learning in a face-to-face environment are also evident in a computer-mediated collaborative learning structure.

Further research into the effect of synchronous computer-mediated communication on solving complex tasks is also necessary. According to Murphy and Collins (1997), research on synchronous computer-mediated communication has been limited to investigations of the recreational use of online chat systems. But the use of these systems for instructional purposes, and specifically for problem-solving tasks, has been explored only through case studies. These case studies support the hypothesis that the benefits of collaborative environment in a face-to-face environment are also found in a computer-mediated environment when students are faced with higher-order cognitive tasks such as problem solving (Hall, 1997; Johnston, 1996; Naidu & Oliver, 1999).

Other research indicates that the quality of interaction between learners in a computer-mediated environment may actually be better than interaction in a face-to-face environment. Findings in a case study suggested that computer-mediated groups seemed to put more thought into the comments they made, thus providing higher quality responses than students who worked face-to-face (Camin, Hall, Quarantillo &

Merenstein, 2001). Hillman (1999) also found that the interaction patterns of computer-mediated groups resembled thoughtful discussions whereas face-to-face interactions resembled recitations. And in yet another study where computer-mediated communications was compared to face-to-face interactions, findings suggest that in the computer-mediated environment there was a tendency to share ideas without the restraints of typical social conventions, which resulted in deeper and more thoughtful discussions (Kruger & Cohen, 1996).

In a recent study that compared computer-mediated collaboration versus individual learning, Uribe, Klein and Sullivan (2003) found that learners who collaborated through a synchronous computer-mediated environment to resolve an ill-defined problem performed significantly better than learners who worked alone. These results support the hypothesis that computer-mediated collaboration has a positive effect on performance when resolving ill-defined problems. The results also showed that, overall, learners preferred to work collaboratively versus individually. Another interesting finding was the high number of on-task interactions (questions, answers and discussions) between members of the dyads, which seems to support the idea that a computer-mediated collaborative environment promotes peer-to-peer communication directly related to the learning task.

Another variable that may influence achievement in a computer-mediated collaborative setting is the size of the collaborative group. Although little empirical exploration exists on this subject, the general rule-of-thumb is that small collaborative groups are better than large groups (Johnson & Johnson, 1988). Lou, Abrami and d'Apollonia (2001) conducted a meta-analysis of 122 studies that compared small group and individual learning with computers and found that groups of two had significantly higher achievement than groups of three to five members. McIsaac and Ralston (1996) observed a distance education undergraduate course and found that collaborative groups larger than five in synchronous environments may require a moderator to be effective.

The size of a computer-mediated collaborative group in synchronous environments is particularly important to consider when solving complex tasks. The most commonly used form of synchronous computer-mediated communication in distance education settings is some form of text-based communication, such as the Internet Relay Chat. An obvious advantage of this type of system is that participants can communicate real-time without the delay experienced in asynchronous systems like email or listservs. A synchronous communication system is also better suited for problem-solving tasks because it more closely resembles a real-world situation (Bastiaens & Martens, 2000). But, according to Murphy and Collins (1997), one of the biggest disadvantages of chat-systems is that turn taking in this synchronous environment can be a problem because there are no visual clues such as body movements, eye-contact, etc., to indicate when someone wants to enter a conversation. In this type of system, all of the users tend to "talk" simultaneously and several conversations may be occurring at the same time, creating a confusing environment. Thus, as the number of members of a synchronous computer-mediated collaborative group increases, the positive effects of collaboration when solving complex tasks may decline. The confusion created by a larger number of participants may have a detrimental effect on the group's ability to solve problems.

The current study examined the effects of two instructional methods (case-based vs. systematic approach) and collaborative group size (teams of two or four) on learner performance in solving ill-defined problems in a web-based environment.

Method

Participants

The participants in this study were 130 cadets from the United States Air Force Academy. All of the participants volunteered to participate in the study. There were 99 male participants (75%) and 31 female participants (25%). The study was incorporated into a three-hour segment of instruction of the cadets' first-year language course. Although the participants' academic ability was not measured, they were expected to be high academic achievers since the Air Force Academy has very high entrance requirements.

Materials

The materials for this study were developed and incorporated into a web-based interface through the Blackboard course management system. Blackboard was used to host the instructional and assessment materials used in the study. Two web-based instructional programs, four assessment scenarios, a scoring rubric and an attitude survey were developed for this study.

Instructional Programs. Two web-based instructional programs were developed to teach students how to solve decision-making problems. One program taught the participants a systematic approach, while the other taught them a case-based approach. A description of each instructional program is presented below.

The web-based program for the systematic approach focused on a four-step problem-solving process derived from the Air Force's "Six-Step Problem-Solving Process" commonly taught to college juniors enrolled in the Air Force ROTC program. This approach was intended to provide students with a tool to solve complex problems. It was developed based on principles for systematic problem solving outlined by Gagné (1985) and Van Merrienboer et al (2002).

The problem-solving process led students through the four steps of (a) problem definition, (b) data gathering, (c) developing and testing possible solutions, and (d) selection of the best possible solution. Within the program, an animated agent taught the steps of the process and used an example scenario showing how to apply each step. Students then applied each step using a practice scenario and received feedback from the animated agent throughout the process. The program prompted students to interact with their team members at several points during instruction.

The web-based program for the case-based approach was designed using the principles outlined by Van Merrienboer et al. (2002) for using case studies to teach problem solving. The program was designed as an open-ended learning environment with some constraints. The same example and practice problem scenarios and animated agent used in the systematic approach were used in this program. However, students were not systematically taught the four steps of the problem-solving process. Instead they were confronted with a given state (problem scenario), a desired state (goal) and a best possible solution as determined by the individual involved in the case (Van Merrienboer et al., 2002). Students were also asked several thought-provoking questions throughout the program, which were intended to trigger deep processing and to develop a better mental model of the problem environment (Van Merrienboer et al., 2002).

Assessment Problem Scenarios. The assessment scenarios used in both web-based programs were modified versions of problem scenarios developed by Hedlund, Sternberg and Psotka (2002) for the Army Research Institute. They were developed from realistic situations encountered by junior Army officers and were modified for the current study to reflect an Air Force theme. Students in all treatment groups were assessed using the same problem scenarios. The following is an example of the types of problem scenarios that were used:

> *"You are a new element leader. Your squadron is preparing to deploy as part of a rapid-response contingency TDY. You assemble your element and tell everyone to start packing equipment in preparation for the deployment that same night. When you come back to inspect their deployment preparation, you find that your airmen have not packed the equipment and are talking to personnel from other elements who are hanging around the area. What should you do?"*

As part of the problem scenario, all students had web access to additional information in the form of simulated interviews with the individuals involved in the problem scenario. Student teams were encouraged to collaborate by having an on-line discussion of the problem scenarios and related information. However, each student was required to provide an individual response the following question for each assessment scenario -- *"What is the BEST POSSIBLE solution for this problem? Make sure to FULLY EXPLAIN the rationale you used to develop your solution."* Students were given four problem scenarios to solve; however, they were told to concentrate on solving the first two scenarios and only do the last two if there was time available.

Scoring Rubric. A scoring rubric was used to evaluate responses to the assessment scenarios. Given the nature of an ill-defined problem scenario, there were a variety of optimal solutions for each problem scenario. Therefore, the rubric was constructed to allow assessment of the validity of the solution proposed by each student, and the quality of the process they used to arrive at the solution. The rubric was designed to evaluate performance in the following four areas: (a) did the student understand what the problem is? (current state or problem definition); (b) did the student understand the desired goal? (end state or goal); (c) did the student follow a logical, well-supported path to achieve a solution? (process); and (d) did the solution produce the desired end-state? Is it the best possible solution given the student's understanding of the problem? (recommended solution). The maximum possible score on the assessment scenarios was 18.

The primary researcher and one other individual trained by the researcher used the scoring rubric to grade student responses to the assessment scenarios. Inter-rater reliability was established by having both raters score the responses for the same ten participants. These were analyzed to determine the correlation between the two raters. The correlation was .94. Once the inter-rater reliability was established, the two raters scored an equal number of responses.

Attitude Survey. A 22-item survey was developed to measure student attitudes. The survey contained 18 Likert-type items with a four-point scale (0=strongly disagree, 3=strongly agree), three open-ended questions, and one forced-choice question.

The18 Likert-type questions on the survey were divided into six different categories with three questions each. However, a factor analysis performed on these 18 questions yielded the following five factors: 1) working with others, 2) collaborating on-line, 3) web-based instructional materials, 4) time available for the program, and 5) transfer of the information learned to other settings. The procedure used for factor extraction is discussed in the design and data analysis section below. Coefficient alpha was computed as an internal consistency estimate of reliability for the Likert portion of the survey and found to be .83.

The open-ended questions asked students to write down the process they used to solve the problems, what they liked best about the program, and what they would do to improve the program. The forced-choice item asked students to select their preference for working on complex problems from the following four choices: alone, with one partner, with two partners, or with three or more partners.

Procedures

One week prior to the study, the primary researcher recruited volunteer participants by reading a script that explained the purpose of the study and described that the top-performing teams would receive a free dinner. The primary researcher conducted the briefing in civilian clothes to preclude the possibility of subtle coercion of the cadets. Once the cadets had made a decision, those who decided to participate were given a copy of the “Subject’s Bill of Rights” and were asked to read and sign an Informed Consent Document.

Immediately following the recruitment portion of the briefing, cadets who volunteered were given approximately 20 minutes of instruction on login procedures for the Blackboard system, how to use the Blackboard system to collaborate, and general instructions on how to navigate throughout the program. Any technical issues related to access to the Blackboard system were clarified during this briefing. Finally, the participants were given a problem scenario and were asked to write-down the process they would use to solve the problem. This provided information on whether the participants had previously learned a process for solving problems.

Participants were randomly assigned to one of four treatment groups - (1) Systematic Approach - Dyads, (2) Systematic Approach - Quads, (3) Case-Based Approach - Dyads, and (4) Case-Based Approach - Quads. On the first day of the study, participants were directed to pre-arranged locations in the computer laboratory, which ensured that each member of a team was physically separated from his or her partners. This procedure was carried out to prevent verbal or bodily communications between the members of the teams, thus simulating a distance-education environment. The participants were verbally instructed to sign-on to the Blackboard system using their assigned username and password and to navigate to the study’s web page to start the program.

The communication between members of the teams took place using the virtual classroom feature of the Blackboard system. This feature allowed students to chat with their partners by entering a virtual classroom that the researcher set up for each team prior to the start of the study. There were no interactions between the participants and the researcher and/or instructors during the study except to remedy any technical difficulties.

Design and Data Analysis

This study used a 2 X 2 posttest only, factorial design. The independent variables were instructional method (case-based vs. systematic approach) and collaborative group size (dyads vs. quads). The primary dependent variable was student performance in solving ill-defined problem scenarios. Time on learning task, time on problem-solving task, and learner attitudes were also analyzed. An analysis of the interactions between the members of the collaborative teams was also conducted.

A 2 x 2 analysis of variance (ANOVA) were conducted on participants' performance on the assessment scenarios, time on learning task, and time spent solving problems with instructional method and group size as the independent variables. Due to a problem with the server hosting the Blackboard program, some data for performance and time were lost resulting in uneven cell sizes. In addition, four participants whose times exceeded three standard deviations from the mean were considered outliers and were excluded from the time-on-task analyses.

A factor analysis was conducted to determine underlying themes on the attitude survey using procedures outlined by Green, Salkind & Akey (2000). The number of underlying factors was determined using all items with eigenvalues greater than one. The factors were then rotated using the VARIMAX method in order to interpret the results. An average score for each factor was computed and separate 2 x 2 ANOVAs were conducted on each factor. Chi-square tests were conducted on the forced choice attitude item. A two-way contingency table analysis was used to determine if a significant relationship existed between treatment group and the preferences expressed by the participants.

Finally, the interactions between members of the teams were automatically recorded by the Blackboard system and analyzed. These data were then categorized by type of interaction. An entry was classified as a question, answer, discussion, encouragement, or off-task behavior. Previous researchers have used similar categories in studies examining small group interactions (Klein & Doran, 1999; Klein & Pridemore, 1994; Uribe, et al, 2003). A percentage for each category was computed and a two-way contingency table analysis was conducted on the data to evaluate if a significant relationship existed between type of interaction and treatment group. Follow-up pairwise comparisons were conducted to evaluate the differences between treatment groups. The Holm's sequential Bonferroni method was used to control for Type I error.

Results

Performance

The first research question investigated the effect of instructional method and group size on performance solving ill-defined problems. Table 1 shows the mean scores and standard deviations for performance on the assessment scenarios. These data reveal that participants averaged 8.12 out 18 points on the assessment scenarios. Participants who worked in dyads had an average score of 8.68 while those who worked in groups of four had an average of 7.5. Participants in the systematic approach treatment had an average score of 8.34 and those in the case-based treatment had an average of 7.92. A 2 x 2 ANOVA conducted on these data revealed that participants working in dyads had significantly higher scores than those working in quads, $F(1,111) = 4.79$, $p = .03$, partial $\eta^2 = .04$.

Table 1. *Mean scores and standard deviations for performance on the problem scenarios.*

Group Size		Instructional Method		Overall
		Systematic	Case-Based	
Dyads	Mean	9.1	8.39	8.68
	(SD)	(2.77)	(3.12)	(2.94)
	n	26	34	60
Quads	Mean	7.66	7.33	7.5
	(SD)	(3.15)	(2.99)	(3.05)
	n	28	27	55
Overall	Mean	8.34	7.92	8.12
	(SD)	(2.99)	(3.08)	(3.04)
	n	54	61	115

Note. The maximum number of points on the problem scenario was 18

Time spent on instruction

The next research questions pertained to the effect of instructional method and group size on time spent on instruction. Participants spent an overall average of 31.3 minutes on the tutorial. Participants who worked in dyads spent an average of 29.2 minutes while those who worked in quads spent an average of 33.7 minutes. Participants who used the systematic approach spent an average of 28.9 minutes on the tutorial and those who worked through the case-based tutorial spent an average of 33.4 minutes. A 2 x 2

ANOVA indicated that participants who worked through the case-based tutorial spent significantly more time-on-task than participants who worked through the systematic approach tutorial, $\underline{F}(1,110) = 3.87$, $p = .05$, partial $\eta^2 = .03$.

Time spent solving problems

The third research question pertained to the effect of instructional method and group size on time spent solving problems. Participants spent an overall average of 18.7 minutes solving the problems. Those who worked in dyads spent an average of 17.4 minutes on the problem scenarios while those who worked in quads spent an average of 20.2 minutes. Participants who used the systematic approach spent an average of 22.1 minutes to solve the problems and those who used the case-based approach spent an average of 16.7 minutes.

A 2 x 2 ANOVA indicated that participants who worked in quads spent significantly more time solving problems than those who worked in dyads, $\underline{F}(1,74) = 7.69$, $\underline{p} = .007$, partial $\eta^2 = .09$. ANOVA also indicated that participants who used the systematic approach spent significantly more time on the scenarios than those who used the case-based approach, $\underline{F}(1,74) = 24.9$, $\underline{p} < .001$, partial $\eta^2 = .25$. The data also showed a significant interaction between group size and instructional method, $\underline{F}(1,74) = 6.6$, $\underline{p} = .01$, partial $\eta^2 = .08$. Follow-up analyses were conducted to evaluate the two simple main effects for instructional method and collaborative group size. To control for Type I error, we set alpha for each test at .025. The results revealed a significant difference between participants who worked in dyads ($\underline{M}$ =19.1) and those who worked in quads (M =24.8) using the systematic approach. The analysis also showed that for participants who worked in quads, there was a significant difference between participants who used the systematic approach ($\underline{M}$ =24.8) and those who used the case-based approach ($\underline{M}$ =16.8).

Participant Attitudes

The next research question investigated the effect of instructional method and group size on participant attitudes. Factor analysis of the 18 Likert survey items revealed five factors or underlying themes. The five identified factors were attitudes toward: 1) working with others, 2) collaborating on-line, 3) instructional materials, 4) time available for the program, and 5) transfer of the information learned to other settings. Item number 16 in the survey ("The scenarios in this program were realistic and applicable to a future Air Force officer") did not correlate with any other item in the survey and it was not included in any of the five factors mentioned above. However, it is important to note that this item was rated highest in the survey with an average rating of 2.61.

The data reveal that participants rated factor number one (working with others) highest ($\underline{M} = 2.22$) and factor two (collaborating on-line) lowest ($\underline{M} = 1.68$). A 2 x 2 ANOVA of each factor revealed mixed results. The data showed a significant interaction for factor two (collaborating on-line), $\underline{F}(1,107) = 7.02$, $p = .009$, partial $\eta^2 = .06$. Follow-up simple main effects analysis revealed a significant difference between participants who worked in dyads (M =1.44, SD = .76) and those who worked in quads (M =1.89, SD = .75) using the systematic approach, $F(1, 107) = 5.1$, $p = .025$. Factor four (time available for the program) indicated a significant main effect by group size, $F(1,107) = 5.09$, $p = .026$, partial $\eta^2 = .05$. The results did not reveal significant results for the other factors in the survey.

Participant Preferences. The attitude survey also included one forced-choice question that asked the participants to choose a preference when solving complex problems. The choices were - by myself, with one partner, with two partners, with more than two partners. A two-way contingency analysis of these data indicated a non-significant relationship between treatment group and preference, $\chi^2(9, \underline{N}=110) = 4.6$, $\underline{p} = .87$, *Cramer's V* = 0.12. When the preference data is analyzed for all participants regardless of their treatment group, the results show that of the 110 participants who answered this question, 12 (11%) indicated a preference to work alone, 39 (35%) indicated a preference to work with one partner, 36 (33%) prefer to work with two partners, and 23 (21%) indicated they prefer to work with more than two partners when solving complex problems, $\chi^2(3, \underline{N}=110) = 17.6$, $\underline{p} = .001$, *effect size* = +0.16.

Problem-solving process used by the participants. Prior to the treatment the participants were given a simple problem scenario and were asked to state the process they would use to solve it. The responses indicated that 91% (118) of the participants did not know a process or heuristic to solve ill-defined problems. After the treatment, the participants were asked to state the process they used to solve the problem scenarios. When the responses to this question are analyzed by instructional method, the data show that 51 out of 60 participants who learned the systematic approach used the approach they learned in the tutorial to solve the problems. Of the 60 participants in the case-based approach who responded to this question, 39 indicated they used a step-by-step process to solve the problems.

Open-ended questions about the program. The top response for what participants liked best in the systematic approach tutorial was the realism of the assessment scenarios, while the participants in the case-based approach tutorial felt working with others was the best thing about the program. Participants in both instructional methods felt changes such as making the assessment scenarios more complex, adding additional practice and feedback, and improving the navigation throughout the tutorial were the best way to improve the program.

Learner-to-Learner Interactions

The last research question dealt with the effect of instructional method and group size on the type and quantity of communications between team members. There were a total of 3,665 communications between members of the groups. These communications were categorized into one of five possible categories: 1) questions, 2) answers, 3) discussions, 4) encouragement, and 5) off-task. The results indicate that 17% were "questions", 16% were "answers", 59% were "discussions", 3% were “encouragement” and 5% were "off-task" entries.

A two-way contingency analysis of these data revealed that treatment group and communication category were significantly related, $\chi^2(12, \underline{N}=3665) = 62.8$, $\underline{p} < .001$, *Cramer's V* = 0.08. Follow-up pairwise comparisons were conducted to evaluate the differences among the four treatment groups. The Holm’s sequential Bonferroni method (Green, et al., 2000) was used to control for Type I error at the .05 level across all six comparisons. Three comparisons (system-dyads vs. case-dyads, system-quads vs.case-dyads, and case-quads vs. case-dyads) of the total number of communications were found to be significant using this method.

Discussion

Performance by Instructional Method and Group Size

The purpose of this study was to investigate the effects of instructional method and computer mediated collaborative group size on learner performance in solving ill-defined problems. The results show that participants who worked in computer-mediated collaborative dyads performed significantly better than did participants who worked in computer-mediated collaborative quads. This finding supports previous research that showed dyads perform better than larger groups (Johnson & Johnson, 1988; Lou, et al., 2001; Collins & Onweugbuzie, 2000).

The fact that dyads interacted more than quads throughout the study may have contributed to this result. Although all participants were able to discuss the problem and related information with others, the results seem to suggest that participants who worked in dyads extracted more benefit from these learner-to-learner interactions than participants who worked in quads.

The finding that a higher level of collaboration results in better performance is supported by findings from other studies where participants worked in a collaborative learning environment (Chang & Smith, 1991; Flynn & Klein, 2001; Johnson & Chung, 1999; Johnson et al., 1991; Mergendoller et al., 1999; Uribe, et al., 2003). The amount of communication between members of dyads when compared to those of quads seems to support the hypothesis that more communication between team members resulted in better performance. The analysis of the team interactions showed that dyads had a higher number of communications than participants who worked in teams of four. This higher number of communications may have led to the generation of a higher number of possible solutions, and ultimately to better performance.

The quality of learner-to-learner interaction may have also led to better performance by participants who worked in dyads over those who worked in quads. An informal qualitative analysis of the interaction data indicates that participants who worked in dyads had higher-quality interactions than those who worked in quads. The exchange of ideas and information between members of dyads appeared to probe deeper into the problem, thus generating better possible solutions; whereas the interactions between members of quads seemed to be superficial, lacking depth and insightfulness. The examples below show a sample interaction for a quad and a dyad and are representative of the interactions observed during the study.

Quad example:

Student 1 > we are doing #1 right?
Student 2 > yeah
Student 3 > right
Student 4 > si
Student 1 > good, i [sic] am ready
Student 4 > me too
Student 2 > let's do it then
Student 2 > what's our solution?
Student 2 > take away the other flight and remind the flight that we have to get on board?
Student 3 > Bust there [sic] chops.
Student 1 > make them pack,
Student 1 > that too,
Student 4 > make them pack
Student 1 > get rid of the other guys

Dyad example:

Student 1 > ok, I'm starting on the first one
Student 2 > ME TOO
Student 1 > what do you think?
Student 2 > i think that b-flight needs to get out of there
Student 2 > but, in the scenario, we're only an [sic] lt and the b-flight com is a captain
Student 2 > so that could pose a problem
Student 1 > good point
Student 1 > but I think the mission is more important so we should have them pack and then they can chat with bflight
Student 2 > i agree
Student 1 > is that our solution then?
Student 2 > well, how do we get them to pack
Student 1 > we order them to
Student 2 > right, but they will still be distracted by b flight
Student 2 > so, i think we should order our troops to pack (direct order)
Student 1 > we need an incentive to pack
Student 2 > and make it very clear to the captain that we want his flight out of there until we're done
Student 1 > ok

While group size had a significant effect on performance, instructional method did not. There are two possible reasons for this result. First, the participants who worked through the case-based approach appeared to have acquired a heuristic that they were able to apply to the problem scenarios. The results show that participants in both instructional methods used a process to solve the problems. Most participants in the systematic approach group indicated they used the 4-step process they learned in the tutorial, while about 65% of the participants in the case-based approach used a three-step heuristic similar to the one learned by the participants in the systematic approach. This finding supports the notion that when learners are faced with a problem situation, they develop a mental model that can then be applied to similar situations (VanMerrienboer, et al., 2002) The fact that both treatment groups applied a step-by-step approach to solve the problems may have equalized their performance on the assessment scenarios.

Another reason for the non-significant difference between instructional methods may have been the amount of time spent on the program. The survey results showed that students felt they could have benefited from additional practice, additional examples and, perhaps most importantly, additional time. The limited amount of spent on instruction may also have led to the overall poor performance by both treatment groups on the problem scenarios. Participants spent an overall average of 31.3 minutes on the tutorial and 18.7 minutes solving the problems. Although poor performance on problem-solving tasks has been evident in previous studies (Flynn & Klein, 2001; Uribe, et al, 2003), it is possible that a longer treatment could have resulted in higher overall scores.

Time Spent on Instruction

The results for time spent on instruction revealed that participants who used the case-based tutorial spent significantly more time on instruction than participants who used the systematic approach. Case-based participants may have spent more time on the learning task because of key differences in the two instructional programs. In the systematic approach participants were taught each step of a problem-solving process, and then were directed to use each step in a practice scenario. Conversely, the case-based approach program was designed to allow the participants more opportunity to reflect on the problem scenarios and to formulate alternative paths to a best possible solution. This difference in the teaching/learning approach may have had an impact on the time spent on the tutorial.

Another explanation may be that the participants in the case-based approach spent more time on instruction because of the increased cognitive load of developing a heuristic or mental model that could be applied to other similar problems (Van Merrienboer, 2002). Participants who learned the systematic approach learned a heuristic directly from the instruction, and therefore did not have the additional mental burden of formulating a step-by-step approach.

Time Spent Solving Problems

In contrast to the results for time spent on instruction, participants who used the systematic approach spent more time solving the problems than did participants in the case-based approach. Participants who used the systematic approach may have spent more time methodically applying the steps of the problem-solving process to solve the scenarios. Although participants in the case-based approach also used a heuristic to solve the problems, the steps appear to have less structured than those used by the systematic approach participants. This may have contributed to the difference in the time spent solving problems.

In addition to the significant difference between participants in the systematic approach and the case-based approach, the results for time spent solving problems exhibited a significant interaction between instructional method and group size. The data show that quads that used the systematic approach spent significantly more time on the problems than did quads using the case-based approach. However, there was no significant difference between dyads who used the systematic approach and those who used the case-based approach. Quads that used the systematic approach may have spent more time on the problem because applying a more structured process to solve the problems required additional coordination than applying a more open-ended case-based approach.

In addition to the higher degree of coordination required to apply a systematic approach to solve problems, confusion arising from synchronous computer-mediated communication in large groups might have also contributed to these results. An informal qualitative analysis of the data indicates that participants who worked in quads experienced some confusion during the communication process. The larger number of team members in the quads might have been a factor in creating a confusing learning environment. This is in line with other research that has shown larger groups may require a moderator to be effective (McIsaac and Ralston, 1996). The attempt to apply a more structured approach in this somewhat disorganized communication environment may have led quads to spend more time solving the problems than participants working in dyads. The example below shows a typical exchange between four members of a quad where the conversation seems somewhat disjointed:

Student 1 > hey guys
Student 2 > yeah
Student 2 > sounds good
Student 1 > I just read that thingy
Student 3 > what thingy
Student 2 > are we "ready to answer the questions"?
Student 1> sure

Student 4 > yes
Student 3 > you guys done with that
Student 2 > with what?
Student 3 > with those answers
Student 2 > not yet
Student 4 > yes

Participant attitudes

In general, participants indicated a preference for working collaboratively. The data specifically show that a majority of participants prefer to work with one or two partners when solving ill-defined problems. Other research has also shown a preference by learners to work in collaborative environments (Flynn & Klein, 2001; Johnson et al., 1991; Lou et al., 2001, Uribe, et al., 2003). However, collaborating on-line was rated lowest by the participants, indicating that they did not necessarily enjoy collaborating using a text-based system. This supports other research, which found that on-line collaboration was rated low by participants due to the difficulties of communicating via a computer (Uribe, et al., 2003).

However, participant attitudes toward collaborating on-line were not constant across the instructional methods. Participants who worked in quads and used the systematic approach liked collaborating on-line significantly more than dyads who used the systematic approach; whereas, there was not a significant difference between dyads and quads who used the case-based approach. Since dyads appear to have attempted more in-depth discussions of the problems, they may have felt more frustrated attempting to collaborate in a more structured environment. However, participants who worked in quads did not experience the same level of frustration using the systematic approach because their communications remained at a superficial level.

The survey results also show that the participants enjoyed the web-based learning materials. The instructional materials were listed as one of the top three things participants liked best in both instructional methods. This finding supports other research where student attitudes toward web-based instruction have been found to be positive (Adelskold, 1999; McIsaac & Ralston, 1996; Savenye, 2001; Uribe, et al., 2003).

The survey data also indicated that most participants felt that what they had learned in this study would transfer to other situations. This is in line with similar studies that have found a positive attitude towards transfer of problem-solving skills to similar situations (Uribe, et al., 2003). Participants felt they would use this process in their professional lives because the problems they faced in the study were directly related to their future career as Air Force officers.

Learner-to-Learner Interactions

The analysis of the learner-to-learner interactions showed that the number of communications in each category varied significantly as a function of treatment group. Results revealed that participants who worked in case-based dyads had significantly more communications than the other treatment groups. These data also showed that dyads collaborated more than quads in both the systematic and case-based approaches. This additional collaboration might have been a factor in dyads performing better than quads. This finding is in line with the results found by Uribe, et al. (2003), where increased collaboration directly contributed to better performance when solving ill-defined problems.

The majority of communications that took place between team members were related to the task. These results also show that over half of the communications were categorized as “discussion,” indicating that the participants in all treatment groups were actively engaged in discussion of the problem or related information. This finding supports other research where a collaborative learning environment produced a high percentage of on-task interactions (Flynn & Klein, 2001; Klein & Doran, 1999; Sherman & Klein, 1995; Uribe, et al., 2003).

Implications

This study has clear implications for the design and delivery of distance learning. The study indicates that grouping learners into dyads instead of quads maybe a better strategy when learning and applying problem-solving skills in a web-based environment. When time-on-task data and number of problems completed is considered in conjunction with performance and group interaction data, findings seem to suggest that dyads are a more efficient grouping than quads. The results show that dyads communicated more, spent less time-on-task, and performed better than quads on the problem scenarios.

The study also suggests that a computer-mediated environment is conducive to collaboration. The high percentage of on-task interactions indicates that computer-mediated collaborative learning should be considered when high-interactivity and exchange of ideas is a major component of a distance learning program. But while learners appear to enjoy working together when solving complex problems, instructional designers and instructors should keep in mind that collaborating using text-based communications appears to have some drawbacks. Additional time should be factored into any program to allow the learner ample time for communication.

The findings of this study also indicate that the systematic approach and the case-based approach may be equally effective for solving ill-defined problems. Some of the evidence indicates that the case-based method may be a better approach if a goal of the learning program is for the learner to develop his or her own rules-of-thumb or heuristics for problem solving. But regardless of method used, the findings clearly suggest that learners enjoy scenarios that closely match the "real world." If possible, problem scenarios representative of the learner's future profession should be used when teaching how to solve ill-defined problems.

Further Research

The results of this study suggest other areas that could be addressed by additional research. Different computer-mediated collaborative group sizes should be explored to determine an optimal size. This study seems to indicate that as group size increases, the efficiency of the group decreases. Research into other groupings could yield data that could be used to confirm this finding and to develop a prediction model for the effect of group size on performance when solving ill-defined problems.

The instructional methods used to teach problem solving could also benefit from additional research. Variations of the systematic and case-based strategies for problem solving should be explored to determine the best approach for specific problem-solving tasks. For example, the ill-defined problems identified by Jonassen (2000) such as decision-making, strategic performance, case analysis and design problems could be examined. Continued research on the different facets of computer-mediated collaboration should help us identify the most effective instructional practices to effectively promote the learning of problem-solving skills in a computer-mediated, distance learning environmen

References

Adelskold, G. (1999). Problem-based distance learning of energy issues via computer network. Distance Education*, 20 (1),*129-143.

Ahern, T.C., & El-Hindi, A.E. (2000). Improving the Instructional Congruency of a Computer Mediated Small Group Discussion: A Case Study in Design and Delivery. Journal of Research on Computing in Education*, 32(3),* 2000. Online. Available at: *http://ehostvgw21.epnet.com/ehost.asp?key=204.179.122.140_8000_-1288567452&site=ehost&return=n&profile=web*, November 24, 2001

Alavi, M. (1994). Computer-Mediated Collaborative Learning: An Empirical Evaluation. MIS Quarterly, June, 159-174.

Bastiaens, T.J., & Martens, R.L. (2000). Conditions for Web-Based Learning with Real Events. In B. Abbey (Ed.), Instructional and Cognitive Impacts of Web-Based Education.(pp. 2-31). London UK: Idea Group Publishing.

Benjamin, S., & Hamdy, H. (1993). Physicians know how to be good educators too: A comparison of the patient problem solving process and the instructional design process. Medical Teacher*, 15(2/3),* 175-179.

Bligh, J. (1995). Problem-based learning in medicine: an introduction. Post-Graduate Media Journal*, 71(8)*, 323-326.

Brandon,D & Hollingshead, A.B. (1999). Collaborative Learning and Computer-Supported Groups. Communication-Education*, 48(2),* 109-126.

Brush, T., & Uden, L. (2000) Using Computer Mediated Communications to Enhance Instructional Design Classes: A Case Study. International Journal of Instructional Media*, 27(2)*, p. 157. Online. Available at: *http://ehostvgw21.epnet.com/ehost.asp?key=204.179.122.140_8000_-1288567452&site=ehost&return=n&profile=web*, November 24, 2001

Camin, C., Glicken, A., Hall, M., Quarantillo, B., & Merenstein, G. (2001). Evaluation of Electronic Discussion Groups as a Teaching/Learning Strategy in an Evidence-Based Medicine Course: A Pilot Study. Education for Health: Change in Learning and Practice*, 14(1),* 21.

Chang, K.Y.R & Smith, F. (1991). Cooperative Learning and CALL/IVD in Beginning Spanish:

An experiment. *Modern Language Journal, 75(2),* 205-211.

Collins, K.M., & Onweugbuzie, A.J. (2000, November). Treatment By Aptitude Interactions as a Mediator of Group Performance in Research Methodology Courses. Paper presented at the annual meeting of the Mid-South Educational Research Association, Lexington Kentucky.

Damon, W. & Phelps, E. (1988). Strategies uses of peer learning in children's education. In T. Beradt & G. Ladd (Eds.), Children's peer relations. New York: Wiley.

Doran, M.S. (1994). The effects of individual, cooperative and collaborative learning structures using a computer simulation in accounting. Unpublished doctoral dissertation, Arizona State University, Tempe.

Edens, K.M. (2000). Promoting Communication, Inquiry, and Reflection in an Early Practicum Experience via an on-line Discussion Group. Action in Teacher Education, 22(2A), 14-23.

Flynn, A.E. & Klein, J.D. The Influence of Discussion Groups in a Case-Based Learning Environment. Educational Technology Research and Development, 49(3*),* 71-86.

Gagné, R.M. (1985). The Conditions of Learning and Theory of Instruction (4th Ed*).* New York: Holt, Rinehart and Winston.

Green, S.B., Salkind, N.J., & Akey, T.M. (2000). Using SPSS for Windows: Analyzing and Understanding Data. (2nd Ed). Upper Saddle River, New Jersey: Prentice Hall.

Hall, D. (1997). Computer mediated communication in post-compulsory teacher education. Open Learning, *12 (3)*, 54-56.

Hedlund, J., Sternberg, R.J., & Psotka, J. (2002). Identifying the abilities involved in the acquisition of tacit knowledge. Manuscript in preparation.

Hillman, D. C. A. (1999). A new method for analyzing patterns of interaction. The American Journal of Distance Education *,*13 (2), 37-47.

Hooper, S. (1992). Effects of peer interaction during computer-based mathematics instruction. Journal of Educational Research, 85(3), 180-189.

Hooper, S. (1994). Cooperative Learning and Computer-Based Instruction. Educational Technology Research and Development, 40(3), 21-38.

Hooper, S., & Hannafin, M.J. (1991). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. Educational Technology Research and Development, 39(3), 27-40.

Johnson, S.D. & Chung, S.P. (1999). The Effect of Thinking Aloud Pair Problem Solving (TAPPS) on the troubleshooting ability of Aviation Technician Students. Journal of Industrial Teacher Education, 37(1), 7-25.

Johnson, D.W., & Johnson, R.T. (1988). Cooperation in the classroom. Edina, MN: Interaction Book Company.

Johnson, D.W., Johnson, R.T., & Holubec, E.J. (1990). *Circles* of learing: Cooperation in the classroom. Edina, MN: Interaction Book Company.

Johnson, D.W., Johnson, R.T., & Smith, C.A. (1991). Cooperative Learning: Increasing College Faculty Instructional Productivity, Washington, D.C.: School of Education and Human Development, The George Washington University,

Johnston, V.L. (1996). Toward a Global Classroom Using Computer-Mediated Communications at UAA. University of Alaska Anchorage. Vocational Teacher Education Research Report (Eric document # ED356759)

Jonassen, D.H. (2000). Toward a Design Theory of Problem Solving. Educational Technology Research and Development, 48(4), 63-85.

Jonassen, D.H., Davidson, M., Collins, M., Campbell, J., & Bannan Haag, B. (1995). Constructivism and Computer-Mediated Communication in Distance Education. The American Journal of Distance Education, 9(2), 7-26.

Jonassen, D.H., Dyer, D., Peters, K., Robinson, T., Harvey, D., King, M. & Loughner, P. (1997). Cognitive flexibility hypertexts on the Web: Engaging learners in meaning making. In B.H. Khan (Ed.), Web-based Instruction (pp. 361-368), Englewood Cliffs, NJ: Educational Technology Publications Inc.

Jonassen, D., T. Prevish, D. Christy, & E. Stavrulaki. (1999). Learning to solve problems on the Web: aggregate planning in a business management course. Distance Education*, 20 (1)*. 49-63.

Julian, M., Kinzie, M.B., & Larsen, V.A. (2000). Compelling case experiences: Performance, practice, and application for emerging instructional designers. Performance Improvement Quarterly, 13(3), 164-201.

Kinzie, M.B., Hrabe, M.B., & Larsen, V.A. (1998). An Instructional Design Case Event: Exploring issues in Professional Practice. Educational Technology Research and Development, 46(1), 53-71.

Klein, J.D. & Doran, M.S. (1999). Implementing Individual and Small Group Learning Structures with a Computer Simulation. Educational Technology Research and Development, 47(1), 97-110.

Klein, J.D. & Pridemore, D.R. (1994). Effects of orienting activities and practice on achievement, continuing motivation, and student behaviors in a cooperative learning environment. Educational Technology Research and Development, 42(4), 41-54.

Kruger, L.J., & Cohen, S. (1996). Using the Internet to Extend Training in Team Problem Solving. Behavior Research Methods, Instruments and Computers, 28(2), 248-253.

Laffey, J., T. Tupper, D. Musser, & J. Wedman. (1998). A computer-mediated support system for project-based learning. Educational Technology Research and Development, 46 (1), 73-86.

Lave, J. (1988). Cognition in practice: Mind, mathematics and culture in everyday life, Cambridge University Press, Cambridge, UK.

Leflore, D. (2000). Theory Supporting Design Guidelines for Web-based Instruction. In B. Abbey (Ed.), Instructional and Cognitive Impacts of Web-Based Education. (pp. 102-117). London UK, Idea Group Publishing.

Levin, B.B. (1995). Using the case method in teacher education: The role of discussion and experience in teachers thinking about cases. Teaching and Teacher Education, *11(1)*, 63-79.

Lotan, R.A., Cohen, E.G., & Hothuis, N. (1994, April). Talking and Working Together: Conditions for Learning in Complex Instruction. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Maskit, D., & Hertz-Lazarowitz, R. (1986, April). Adults in Cooperative Learning: Effects of Group Size and Group Gender Composition on Group Learning Behaviors (A Summary). Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

McIsaac, M.S. & Ralston, K.D. (1996). Teaching at a Distance Using Computer Conferencing. Tech Trends, 41(6), 48-53.

Mergendoller, J.R., Bellisimo, Y., & Maxwell, N.L. (2000). Comparing Problem-Based Learning and Traditional Instruction in High School Economics. Journal of Educational Research, 93(6), 374-383.

Miller, S.M., & Miller, K.L. (2000). Theoretical and Practical Considerations in the Design of Web-Based Instruction. In B. Abbey (ed.), Instructional and Cognitive Impacts of Web-Based Education.(pp. 156-177). London UK: Idea Group Publishing.

Moore, M.G. & Kearsley, G. (1996). Distance Education: A Systems View. New York: Wadsworth Publishing Company.

Murphy, K.L., & Collins, M.P. (1997). Development of Communication Conventions in Instructional Electronic Chats. Journal of Distance Education, 12(1), 177-200.

Naidu, S. & Oliver, M. (1999). Critical Incident-Based Computer Supported Collaborative Learning. Instructional Science, 27(5), 329-354.

Oliver, R., & Herrington, J. (2000). Using Situated Learning as a Design Strategy for Web-Based Learning. In B. Abbey (ed.), Instructional and Cognitive Impacts of Web-Based Education (pp. 178-191). London UK: Idea Group Publishing.

Phye, G.D. 2001. Problem-Solving Instruction and Problem-Solving Transfer: The Correspondence Issue. Journal of Educational Psychology, 93(3), 571-578.

Shulman, J (Ed.). (1992), Case Method in teacher education. New York: Teachers College Press.

Simonson, M., Smaldino, S., Albright, M., & Zvacek, S. (2000). Teaching and Learning at a Distance: Foundations of Distance Education. Upper Saddle River, N.J.: Merrill an Imprint of Prentice Hall.

Smith, B.L, & McGregor, J.T. (1992). What is collaborative learning? In A. Goodsell, M. Maher & V. Tinto (Eds.), Collaborative Learning: A sourcebook for higher education (pp. 9-22). University Park, PA: National Center for Postsecondary Teaching, Learning and Assessment.

Spector, M.J. & Davidsen, P.I. (2000). Designing Technology-Enhanced Learning Environments. In B. Abbey (ed.), Instructional and Cognitive Impacts of Web-Based Education. (pp. 241-261). London UK, Idea Group Publishing.

Spiro, R.J., Feltovich, P.J., Jacobson, M.J., & Coulson, R.L. (1993). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T.M. Duffy and D.H. Jonassen (Eds.), Constructivism and the technology of instruction: A conversation, (pp. 57-75), Englewood Cliffs, NJ: Lawrence Erlbaum Associates.

Stepich, D.A., Ertmer, P.A., & Lane, M.M. Problem-Solving in a Case-Based Course: Strategies for Facilitating Coached Expertise. Educational Technology Research and Development, 49(3), 53-67.

Stepien, W., & Gallagher, S. (1993). Problem-Based Learning: As Authentic as it gets. Educational Leadership, 50(7), 25-28.

Tillman, B.A. (1995). Reflections on case method teaching. Action in Teacher Education, 17(1), 1-8.

Tyan, N.N. & Hong, F.M. (1998, February). When Western Technology Meets Oriental Culture: Use of Computer-Mediated Communication in a Higher Education Classroom. Paper presented at the Annual Meeting of the Association for Educational Communications and Technology. St Louis, MO.

Uribe, D., Klein, J. & Sullivan, H. (2003). The Effect of Computer-Mediated Collaborative Learning on Solving Ill-defined Problems. Educational Technology Research and Development, 51(1), *5-19*

Van Merrienboer, J.J.G., Clark, R.E., Moore, J.A., & de Crook, M.B.M. (2002). Blueprints for Complex Learning: The 4C/ID Model. Educational Technology Research and Development, 50(2), 39-64.

Van Streun, A. (2000). Representations in Applying Functions. International Journal of Mathematical Education in Science & Technology, 31(5), 703-726.

Vrasidas, C. & McIsaac, M.S. (1999). Factors Influencing Interaction in an Online Course. The American Journal of Distance Education, 13(3), 22-36.

Instructional Effects of Varied Concept Mapping Strategies and Prior Knowledge in Facilitating Achievement of Different Educational Objectives in a Web Based Learning Environment

Charles Xiaoxue Wang
Georgia State University

Francis M. Dwyer
Pennsylvania State University

Abstract

This study examined instructional effects of three concept mapping strategies and prior knowledge in facilitating student achievement in web-based learning environment. The three concept mapping strategies were concept matching, proposition identifying and student generated concept mapping. The prior knowledge levels of the subjects were identified through a general physiology test. The instructional material was a 2000-word expository text about the human heart. Achievement was measured by the identification, terminology, comprehension tests and the total test. The criterion tests measured achievement of different educational objectives at factual, conceptual, and rules and principles levels and general academic performance of university undergraduate students. The subjects were 290 undergraduate students from a large state university. The subjects completed the prior knowledge test, and participated in a 50-minute workshop on concept mapping one week prior to the experimental treatment. The subjects were randomly assigned to one of four treatment groups. Multivariate analysis of variance was used for data analysis. Statistical significant results were found in this study at the .05 alpha level.

Introduction

Gagné and Briggs [1] defined learning as "… the set of cognitive processes that transforms the stimulation from the environment into the several phases of information processing" necessary for acquiring a new capacity (p. 43). How the information is processed during learning influences information retrieval and recall and, consequently, the ability of the learner to apply the information in his/her later use. Ausubel [2] advocated meaningful learning. Meaningful learning, according to Ausubel [2], occurs when a learner consciously links new knowledge to what the learner already knows and makes that piece of information "meaningful" to himself.

Prior knowledge is commonly defined as the ability, knowledge, and skills possessed by the learner before instruction [3]. It is the experience and knowledge a person brings to a learning situation, which greatly influences how the learning material is comprehended, interpreted, and recalled [6]. Many researchers [2, 4, 5, 6, 7] maintained that the prior knowledge is an important factor that influences learning. Some of them [5, 6] provided empirical evidence that prior knowledge plays a significant part in how learners interact with the learning material and perform on various learning tasks as measured by different criterion tests.

Concept maps are most commonly defined as graphic representations of knowledge of a domain. A concept map consists of a set of nodes representing concepts, objects, or actions connected by directional links that define the relationships between and among those nodes [8]. In concept maps, concepts are arranged hierarchically so that the most inclusive, general concepts appear at the top of the map, with less inclusive, subordinate concepts below [3, 9, 10, 11, 12, 13]. Concept mapping in this study is defined as a learning strategy that helps learners to achieve specific learning objectives through the use of or creation of concept maps.

The web-based learning environment refers to those learning situations where instructions are delivered over computer network and displayed by web-browser [14]. In a web-based learning environment, concept maps are widely used (a) as a navigation tool in the form of a site map to help the learners to navigate the course site [15, 16, 17], (b) as a learning tool [15, 16, 18, 19, 20], and (c) as an evaluation tool [16, 20]. This study explores instructional effects of concept mapping and prior knowledge in facilitating achievement of different educational objectives in a web-based learning environment.

Statement of the Problem and Research Questions

Studies of concept mapping as a tool to facilitate learning are extensive and diverse with subjects from kindergarten through college undergraduate education [21]. Research provides empirical evidence that concept mapping is a useful learning strategy [22, 23]. Among the extensive and diverse studies, research on the appropriate use of concept mapping strategies regarding their effectiveness in facilitating the achievement of specific learning objectives and their relationships with learners' prior knowledge level is limited. This study, different from previous studies, identifies the variables in construction of concept maps and purposefully tests the instructional effects of concept mapping in facilitating student achievement when those identified variables are manipulated in a web-based learning environment. Concept map development consists of three important elements: the concepts, propositions and its hierarchical structure. By controlling and manipulating these variables in concept mapping, it is hypothesized that different levels of cognitive processes will be affected and therefore instigate different levels of information processing. This study tried to answers the following research questions

Q1: *Are different types of concept mapping strategies equally effective in facilitating student achievement of different educational objectives measured by the different criterion tests?*

Q2: *Are there any significant differences in achievement among students identified as possessing low, medium, and high prior knowledge receiving different treatments on the different criterion tests?*

Q3: *Is there a significant interaction between levels of prior knowledge and instructional treatment type on the different criterion tests?*

Methods

Subjects, Learning Material, and Criterion Measures

Subjects in this study were 290 undergraduate students from an educational psychology and a statistics class at a large state university. The learning material used in the experiment was a meaningful 2,000-word expository text describing the human heart, including its parts, locations and functions during systolic and diastolic phases. The material was developed by Dwyer [24]. Dwyer identified basic performances that college students over a wide variety of disciplines were generally expected to perform in a course study. He maintained that "…students were expected to: (a) to learn terminology and facts basic to the course content, (b) identify locations and/or positions, (c) construct and/or understand relationships, and (d) engage in problem solving activities" [25]. Based on findings from his research, Dwyer [24] developed the learning material and the criterion tests in order to assess students' achievement on the educational objectives derived from those basic academic performances above. A general test of physiology was used to categorize the subjects into different prior knowledge levels as low and high. The criterion tests in the experiment were the tests of identification, terminology, and comprehension, and a total criterion tests. The following description of the criterion tests was adapted from Dwyer [25].

Physiology Test The prior knowledge test of physiology consisted of 36 multiple-choice question items. The first 30 items were taken from *Reviewing Biology* [26]. The remaining six items were developed by Dwyer [24]. The objective of the prior knowledge test was to determine the level of general knowledge the subjects had about human physiology. The reliability coefficient for physiology test in this study was .79.

Identification Test The objective of the identification test is to evaluate the subjects' ability to identify parts or positions of an object. This 20-item multiple-choice question test measures the subjects' the ability to use visual cues to discriminate one structure of the heart from another and to associate specific parts of the heart with their proper names. The reliability coefficient for this criterion test in this study was .77.

Terminology Test: This test of 20 item multiple-choice questions was designed to measure knowledge of specific facts, terms, definitions, and concepts in the instructional material. The reliability coefficient for this criterion test in this study was .73.

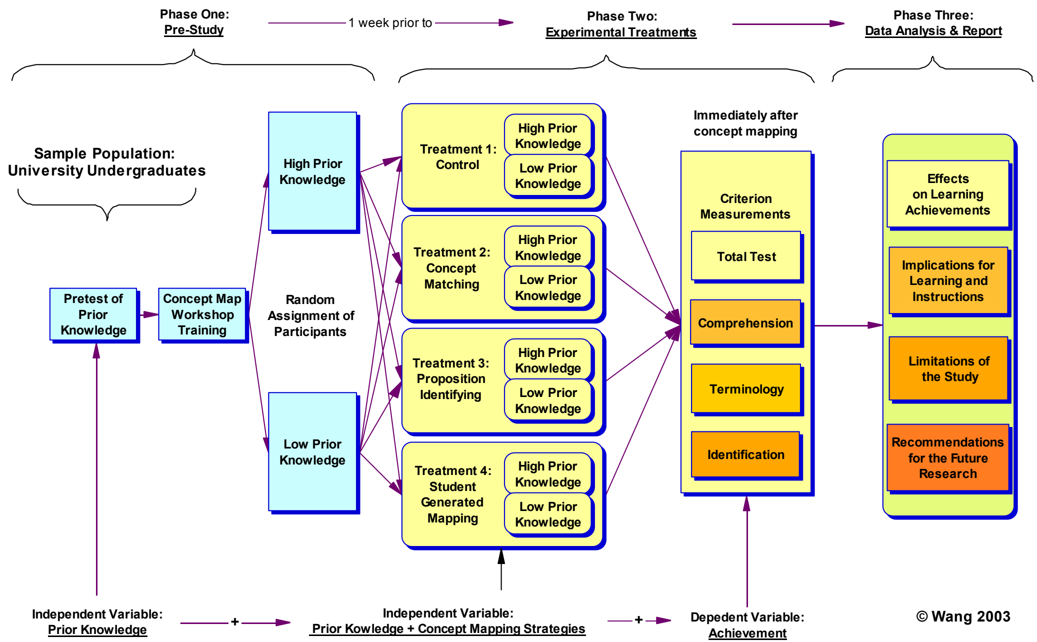
Comprehension Test: This test of 20-item multiple-choice questions was designed to measure the students' ability to comprehend the instructional material content. The reliability coefficient for this criterion test in this study was .70.

The Total Test: The items in the identification, terminology and comprehension tests were combined to give a total criterion score. The score was used to measure students' general academic performance. The reliability coefficient for this criterion test in this study was .88.

Procedures and Treatments

The study consisted of three phases: pre-experiment concept mapping workshop, the experiment, and data analysis and report as illustrated by Figure 1.

Figure 1: Design of the Study



In the pre-experiment phase, a general test of physiology was administered and 50-minute concept mapping workshops were conducted one week prior to the experiment. According to the results of the physiology test, stratified random assignment was used to assign subjects into one of the four experimental groups.

The experiment was conducted at the same time in two university computer labs with Internet access. An experimental information packet for each group was given to the subjects at the door of the computer labs. The packet gave the subjects a study website URL, and explained the rules and regulations for the study. All the instructions and learning material were delivered online at the given websites. The pre-programmed function of the web site collected all study data and the data collected were directly imported into the Statistical Package for the Social Sciences (SPSS) version 11.5 for analysis. The four experimental treatments in this study were Treatment 1: Control, Treatment 2: Concept matching mapping, Treatment 3: Proposition identifying mapping, and Treatment 4: Student-generated mapping.

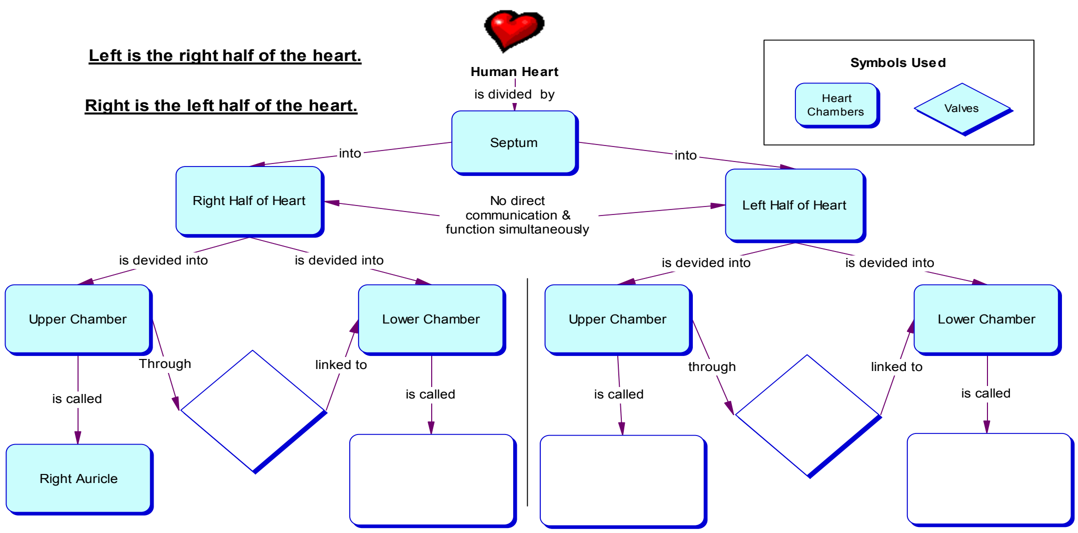
Treatment 1: Control: The subjects in this treatment were required to interact with the learning material at their own pace and then to take the online criterion tests immediately after they finished their study. No concept mapping strategies were used in this treatment.

Treatment 2: Concept Matching: The subjects in this treatment (concept matching) were required to interact with the learning material at their own pace while using the concept matching mapping strategy to help them study the learning material. They took the online criterion tests after they finished their reading of the learning material and concept mapping activity. Figure 2 is an example of the concept matching mapping activity in Treatment 2.

Figure 2: Sample Mapping Activities for Treatment 2

Direction: Please complete the concept map by filling in the concepts according to the learning material **Part 2 & 3**.

Concept Map of Heart Structure

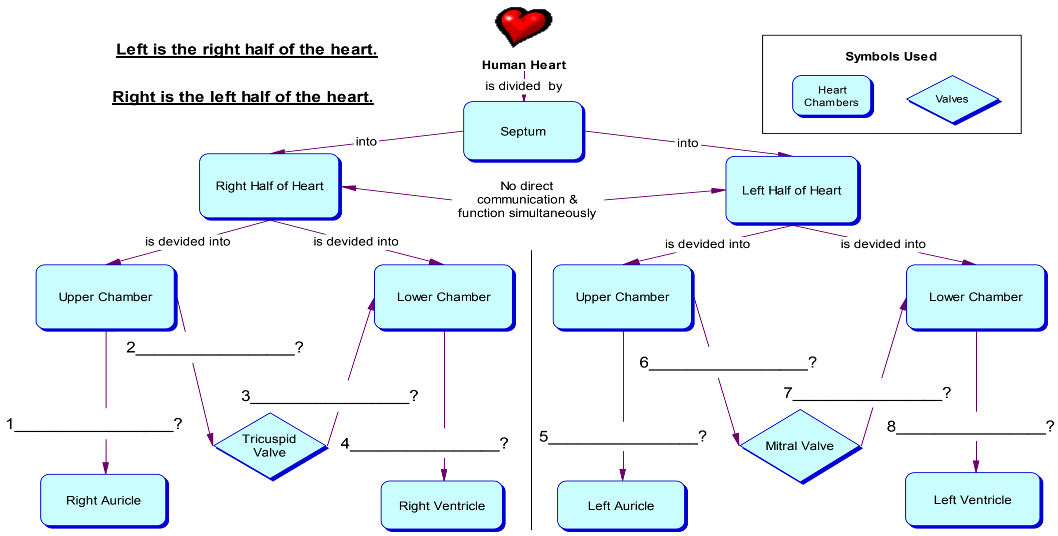


Treatment 3: Proposition Identifying Mapping: The subjects in this treatment were required to interact with the learning material at their own pace while using proposition identifying mapping to help them study the learning material. They took the online criterion tests after they finished study and the concept mapping activity. Figure 3 is an example of the proposition identifying mapping activity in Treatment 3.

Figure 3: Sample Mapping Activities for Treatment 3

Direction: Please complete the concept map by filling in link words according to the learning material **Part 2 & 3**.

Concept Map of Heart Structure



Treatment 4: Student-generated Mapping : The subjects in this treatment were required to interact with the learning material at their own pace and to generate their own concept maps depicting their understanding of the learning material during their study. They took the online criterion tests after they finished the concept mapping activity. Figure 4 is an example of the student-generated concept mapping activity in Treatment 4.

Figure 4: Sample Mapping Activities for Treatment 4

Direction: Please create a concept map about heart structure according to the learning material **Part 2 & 3**.

Concept Map of Heart Structure

The study collected 265 sets of data. However, 34 sets of data collected were incompleted with one or two criterion test data missing. Consequently, 231 sets of data were used in the analysis. Multivariate analysis of variance (MANOVA) was used to test the instructional effects of the three concept mapping strategies and prior knowledge with alpha level set at .05. According to the results of physiology test, subjects are categorized into low prior knowledge group and high prior knowledge. Table 1 reports the descriptive the data of the pre-test of physiology.

Table 1: *Descriptive Data of the Pre-Test of Physiology*

	Prior Knowledge Levels					
	Low Prior Knowledge (0-21)			High Prior Knowledge (22-36)		
	M	SD	n	M	SD	n
Treatment 1	18.16	2.238	31	24.71	2.692	28
Treatment 2	18.77	2.065	26	25.03	3.104	31
Treatment 3	18.97	2.042	30	24.83	2.692	30
Treatment 4	18.74	2.141	27	25.04	2.912	28
Total	18.65	2.120	114	24.91	2.825	117

Treatment 1: Control
Treatment 2: Concept matching mapping
Treatment 3: Proposition identifying mapping
Treatment 4: Student-generated concept mapping
The maximum score for the Test of Prior knowledge is 36.
Total number of subjects in this study is 231.

MANOVA on the achievement indicated that there were significant differences at the .05 level on the criterion tests among all treatment groups and among the prior knowledge levels in this study ($P<.001$, df = 9, Pillais Trace F =3.357). There was no signification interaction found between the types of concept mapping treatments and prior knowledge levels (P=.775, df =9, Pillais Trace F =.735). Table 2 summarizes the results of MANOVA on each criterion test.

Table 2: *MANOVA Results of the Study*

Dependent Variables	Multivariate: F	Group Univariate F df = 3	Prior Knowledge Univariate F df = 1
Treatment Group Effect Pillais Trace	3.357 (p< .001; Ho: df = 9)		
Group * Prior Knowledge Pillais Trace	.671(p = .735; Ho: df = 9)		
Identification Test		4.221 (p= .006)	50.237 (p< .001)
Terminology Test		9.545 (p< .001)	47.777 (p< .001)
Comprehension Test		7.606 (p< .001)	44.514 (p< .001)
Total Test		8.476 (p< .001)	63.237 (p< .001)

Design: Intercept + Group + Prior Knowledge + Group * Prior Knowledge.

The post hoc tests (Bonferroni) located specific significant differences among treatment groups on each criterion test for subjects of different prior knowledge levels. Table 3 reports those significant and insignificant differences found among the low prior knowledge group and high prior knowledge group.

Table 7: *Significant Differences for Prior Knowledge Groups (Post Hoc Test of Bonferroni)*

	Prior Knowledge	Identification	Terminology	Comprehen-sion	Total Test
T 2	HPK	X	T 2 > T 1	T 2 > T 1	T 2 > T 1
	LPK	T 2 > T 1	T 2 > T 1	T 2 > T 1	T 2 > T 1
T 3	HPK	X	X	X	X
	LPK	X	T 3 > T 1	X	T 3 > T 1
T 4	HPK	X	X	X	X
	LPK	X	X	X	X

T 1: Control
T 2: Concept matching mapping
T 3: Proposition identifying mapping
T 4: Student generated concept mapping
HPK: High Prior Knowledge Group
> Significant differences founds between the two groups at .05.
LPK: Low Prior Knowledge Group
X = No significances found between this group and the other groups.

Interpretation and Discussion

The criterion tests in this study measured learning achievement at four different levels as summarized by Table 8. The interpretation and discussion of both significant and insignificant differences on achievement in this study are based on these four levels.

Table 8: *Criterion Measures, Levels of Performance, and Learning Hierarchy*

Criterion Measures	*Levels of Performance*	*Learning Hierarchy*
Total Test	*General academic performance of undergraduates for a course study in a university*	*Overall*
Comprehension	*Comprehending complex procedures and explaining a how a system works*	*Rules & Principles*
Terminology	*Defining terms and rephrasing concepts relating to the learning material*	*Conceptual*
Identification	*Identifying parts and position of an object relating to the learning material*	*Factual*

Data analysis indicated that different concept mapping strategies had different effects in facilitating achievement of different educational objectives for subjects of different prior knowledge levels. For subjects identified as possessing low prior knowledge, concept matching mapping (Treatment 2) is most effective in facilitating learning at the factual levels, conceptual level, rules and principles level and achievement in general academic performance among the three concept mapping strategies (Treatment 2, Treatment 3, and Treatment 4). From an instructional design perspective, it is interesting to note that as a learning strategy, concept matching mapping (Treatment 2) has the tendency to lessen the gap between learners with low prior knowledge and learners with high prior knowledge in learning at factual and conceptual levels and in general academic performance as measured by the identification, terminology and the total criterion tests.

For the subjects identified as possessing high prior knowledge, the study revealed very similar results as those subjects identified as possessing low prior knowledge. Concept matching mapping (Treatment 2) was the most effective mapping strategy in facilitating learning at all levels as measured by the terminology, comprehension and total criterion tests. The significant differences existed not only between the concept matching mapping (Treatment 2) and the control (Treatment 1) on the terminology, comprehension and the total criterion tests but also between proposition identifying mapping (Treatment 3) on the identification and terminology, and total criterion tests.

The instructional effects of concept matching mapping (Treatment 2) was decided by the format of this mapping strategy. Pankratius [27] pointed out that students benefit most from concept maps that reveal the connections between concepts by using labeled words. Novak and Gowin [28] also stated that it is the connectedness between and among individual concepts in concept mapping that facilitates meaningful learning. Concept matching mapping (Treatment 2), by labeling all important propositions among the selected concepts in a map, well framed all the key concepts in the learning material, instigated the meaningful and active interaction, and helped the learners process not only the factual information of those key concepts but also comprehend other dimensions of the content (other than the factual information of those key concepts) in the learning material.

The observed instructional effects of concept matching mapping (Treatment 2) were also in alignment with the concept of learning hierarchy [29]. In explaining the phases in the learning hierarchy, Dwyer [25] stated that the more facts that a person is familiar within a content area, the better prepared the person is to relate and combine those factual information into concepts, and the more concepts a person possesses, the easier it is for the person to form generalizations, rules and principles. Dwyer's explanation of the phases of the learning hierarchy can be borrowed to explain the reason why concept matching mapping (Treatment 2), which focuses on facilitating learning at a factual level, can eventually facilitate learning at the concept, and rules and principles levels for those learners with low prior knowledge.

Insignificances among treatments are also interesting. For proposition identifying mapping (Treatment 3), one of the possible factors might be that the treatment itself was confusing. Concept maps without propositions are not well defined and are ambiguous. Different people might perceive these concepts without propositions in their own ways according to their understanding of the learning material as they are required to identify the propositions among these concepts. According to Stewart [30], there could be "… numerous valid propositions that could be generated to link two nodes" (p. 400). The ambiguity in given maps with deleted propositions as those in Treatment 3 could easily confuse the subjects and might cause negative effect on learning.

Theoretically, concepts, according to Novak [11], are perceived regularity in events or objects, or records of events or objects designated by labeling words. The perceived regularity is defined by the attributes of the concepts, which are framed by the propositions among those concepts. This explains why Novak and Gowin [28] maintain that concept mapping facilitates meaningful learning by illustrating the connectedness between and among individual concepts and concept mapping can serve as a learning tool to help learners organize their cognitive frameworks into more powerful integrated patterns with those labeled propositions. Pankratius [27] stated that students benefit most from maps that reveal the connections between concepts using labeled word. Proposition identifying mapping, removing those critical propositions in the given concept maps, might not be a good concept mapping strategy. So, in devising proposition identifying concept mapping, it is advisable to be extra cautious when removing some of the propositions from the concept maps.

Additionally, the fact that subjects were not familiar with proposition identifying mapping and student-generated mapping might be another explanation for those insignificant differences. For instance, student-generated mapping is a very demanding metacognitive strategy that takes a relatively long time, and requires sustained efforts and fairly good concept mapping skill, if it is meant to be effective in facilitating learning. Okebukola [31] stated that the process of generating a concept map required the subjects to think in multiple directions and to switch back and forth between different levels of abstractions. Student generated concept mapping (Treatment 4) overloads the learners with cognitive demands that it could hardly facilitate meaningfully process the information. Both the concept maps generated by the subjects and the frustrations observed in generating concept maps during this study indicated the lack of concept mapping skills among the subjects. As a result, the student-generated concept mapping has failed to facilitate achievement significantly in this study as expected by the researchers.

Schau and Mattern [32] argued that asking students to draw a totally user-generated map imposes a high cognitive demand to extract meaningful representations of their knowledge. "This demand may have caused the students to focus more heavily on the map than on actually understanding the material and establishing mental connections" [33, p.97]. Snead [34] asserted that a lack of familiarity with the concept mapping strategy and inadequate preparation for the use of the concept mapping strategy is likely a factor for the insignificant differences. The subjects in this study, who received only a 50-minute workshop training on concept mapping prior to the study, were certainly not sufficiently prepared with the concept mapping strategies. Successful use of concept mapping requires practice and a willingness of people to be open to a great deal of trial and error and ambiguity [10]. Concept mapping is most effective when accompanied with comprehensive training, instructor guidance, and long-term practice [35]. Providing extra guidance or scaffoldings for learners in generating their concept maps might enhance the learning in a web-based learning environment.

For subjects identified as possessing high prior knowledge, students generated concept mapping did not affect the learning significantly. This finding agrees with reports from other researchers [6, 36, 37], who found in their studies that learners with high prior knowledge performed well following instruction that was incomplete and less structured, allowing the individual to use self-selected or metacognitive learning strategies. In this study, subjects were not given the freedom of choosing their own learning strategies but were required to use different concept mapping strategies. For those high prior knowledge subjects, it might be that forcing them to use unfamiliar learning strategies impedes their learning.

References

Gagné, R. M., & Briggs, L. J. (1979). *Principles of instructional design* (2^{nd} ed.). New York: Holt, Rinehart & Winston.

Ausubel, D. P. (1968). *Educational psychology : A cognitive view.* New York: Holt, Rinhart and Winston, Inc.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and* *instruction.* Hillsdale, NJ: Erlbaum.

Driscoll, M. P. (2000). *Psychology of learning for instruction.* Boston, MA: Allyn and Bacon.

Dwyer, F. M., & Dwyer, C. A. (1990). Effects of student's prior knowledge and presentation mode on achievement (visual/verbal testing) of different educational objectives*.* Paper presented at the Convention of the Association of Educational Communications and Technology. (ERIC Document Reproduction Service No. ED 323 926).

Elliott, C. (1993). *Effects of prior knowledge and various rehearsal strategies on student achievement of different educational objectives.* Unpublished doctoral dissertation. Pennsylvania State University.

Thompson, R. A. & Zamboanga, B. L. (2003). Prior knowledge and its relevance to student achievement in introduction to psychology. *Teaching of Psychology, 30* (2), 96-101.

Alpert, S. R., & Grueneberg, K. (2000). Concept mapping with multimedia on the web. *Journal of Educational Media and Hypermedia, 9*(4), 313–330.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993) *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge.* Hillsdale, NJ: Erlbaum.

Novak, J. D. (1990).Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27 (10).* 937-949.

Novak, J. D. (1998*). Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Plotnick, E. (1997). Concept mapping: A graphical system for understanding the relationship between concepts*. Eric Digest, EDO-IR-97-05.*

Plotnick, E. (2001). A graphical system for understanding the relationship between concepts*. Teacher Librarian: the journal for school library professionals, 28*(4), 42–44.

Clark, D. E. (2000). *A process model and concept map for web-based teaching and learning in postsecondary art education*. Unpublished Doctoral Dissertation, Florida State University.

Gaines, B. R. & Shaw, M. G. (1995). Concept Mapping on the Web, *Proceedings of WWW4: Fourth International World Wide Web Conference*. Boston, December, 1995.

Kommers, P. & Lanzing, J. (1997). Students' concept mapping for hypermedia design: Navigation through World Wide Web (WWW) space and self-assessment. *Journal of Interactive Learning Research, 8*(3/4), 421-455.

Tan, S. C. (2000). The effects of incorporating concept mapping into computer assisted instruction. *Journal of Educational Computing Research 23* (2), 113-131.

Chang, K. E., Sung, Y. T. & Chen, I. D. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning, 17,* 21-33.

Shortridge, A. M. (2001). *Using concept mapping as an interactive learning tool in web-based distance education.* Unpublished Doctor Dissertation. University of Houston.

Tsai, C. C., Lin, S. S. J., & Yuan, S. M. (2001). Student's use of web-based concept map testing and strategies for learning. *Journal of Computer Assisted Learning* (2001) 17. 72-84.

Helton, C. A. (1994). *Comparative investigation of three study strategies: Concept mapping, marking text, and a combination of concept mapping and marking text.* Unpublished Doctoral Dissertation. The Tennessee State University.

Wang, H. F. (1995). *Effects of format and student completion of concept maps on college students' learning.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Winn, W., & Snyder, D. (1996). Cognitive perspectives in psychology. In D. H. Jonassen (ed.), *Handbook of research for educational communications and technology*. New York: Macmillan LIBRARY Reference USA.

Dwyer, F. M. (1965). *An experimental evaluation of the relative effectiveness of selected visual illustrations in teaching science concepts to college students.* Unpublished doctoral dissertation, The Pennsylvania State University, University Park, Pa.

Dwyer, F. M. (1978). *Strategies for improving visual learning.* State College, PA: Learning Services.

Hall, M. A. (1985). *Reveiwing biology.* New York: Amsco School Publications, Inc.

Pankratius, W. J. (1990). Building and organized knowledge-base: Concept mapping and achievement in secondary school physics. *Journal of Research in Science Teaching, 27*(4), 315–333.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.

Gagné, R. M. (1977). *Conditions of learning* (3rd ed.). New York : Holt, Rinehart & Winston.

Steward, J. (1979). Content and cognitive structure: Critique of assessment and representation techniques used by science educational researchers. *Science Education, 63,* 395–405.

Okebukola, P. K. (1990). Attaining meaningful learning of concepts in genetics and ecology. An examination of potency of the concept mapping technique. *Journal of Research in Science Teaching*, *27*(5), 493–504.

Schau, C., & Mattern, N. (1997). Use of map techniques in teaching statistics courses. *The American Statistician, 51*(2), 171–175.

Taricani, E. M. (2002). *Effects of the level of generativity in concept mapping with knowledge of correct response feedback on learning.* Unpublished doctoral dissertation, Pennsylvania State University, University Park, Pa.

Snead, D. (2000). *Concept mapping and science achievement of middle grade students*. Unpublished doctoral dissertation. University of Kentucky.

Roop, K. M. (2002). *Effects of concept mapping as a learning strategy on certificate practical nursing students' achievement and critical thinking development*. Unpublished doctoral dissertation. Wilmington College, DE.

Davey, B., & Kapinus, B. A. (1985). Prior knowledge and recall of unfamiliar information: Reader and text factors. *Journal of Educational Psychology, 56,* 1–39.

Pintrich, P. R., Cross, D. R., Kozma, R. B., & McMeachie, W. J. (1986). Instructional psychology. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology* (pp. 611–651). Palo Alto, CA: Annual Review Inc.

Vicarious Learning Experiences and Goal Setting: Impact on Preservice Teachers' Self-Efficacy for Technology Integration

Ling Wang
Peggy A. Ertmer
Nova Southeastern University

Abstract

This study was designed to explore how vicarious learning experiences and goal setting influence preservice teachers' self-efficacy for integrating technology into the classroom. Two hundred and eighty undergraduate students, enrolled in an introductory educational technology course at a large Mid-western university, participated. Students were divided into eighteen lab sections, which were assigned to one of four conditions (3 experimental and 1 control). Pre- and post-surveys were administered to examine participants' self-efficacy beliefs for technology integration. Results showed significant treatment effects for vicarious experiences and goal setting on participants' judgments of self-efficacy for technology integration. A significantly more powerful effect was found when vicarious learning experiences and goal setting were both present compared to when only one of the two factors was present.

Introduction

In an effort to prepare students for the information age, public schools are increasing access to technology tools by installing more hardware and software in schools, connecting classrooms to the Internet, and providing cable and satellite capabilities (Zehr, 1997; 1998). Yet, despite the increased availability and support for classroom computer use, relatively few teachers have fully integrated computers into their teaching (Becker, 2000; Marcinkiewicz, 1996). Teachers' uses of computers are likely to be influenced by multiple factors including the accessibility of hardware and relevant software, the nature of the curriculum, personal capabilities, and external constraints such as time, equipment, and technical support (Albion, 1999). However, according to Ertmer (1999), "even if every first-order [external] barrier were removed, teachers would not automatically use technology to achieve the kind of meaningful outcomes advocated" (p. 51).

There is substantial evidence to suggest that teachers' beliefs in their capacity to work effectively with technology, that is, their computer self-efficacy, may be a significant factor in determining patterns of classroom computer use (Albion, 1999; Oliver & Shapiro, 1993). For example, according to Eachus and Cassidy (1999), "self-efficacy has repeatedly been reported as a major factor in understanding the frequency and success with which individuals use computers" (p. 2). Compeau, Higgins, and Huff (1999) conducted a longitudinal study with 394 subscribers to a periodical over a one-year interval to test the influence of computer self-efficacy beliefs, outcome expectations, affect, and anxiety on computer use. Their findings provided strong confirmation that computer self-efficacy beliefs had a significant positive influence on computer use. Another study, conducted by Albion (1996), investigated student teachers' dispositions toward computers and their uses of computers in primary school classrooms during a final-year practicum. Results suggested that lack of confidence for teaching with computers was an important factor influencing the levels of computer use by student teachers. Taken together, these studies suggest that teachers' beliefs and, in particular, self-efficacy beliefs, are useful indicators of levels of technology integration. Certainly, they provide sufficient reason to undertake further investigations in this area and to consider approaches to teacher education and professional development that might be effective in increasing self-efficacy for teaching with technology.

Bandura (1986) identified four sources of information used to judge self-efficacy: successful performance attainment; observing the performances of others (vicarious learning); verbal persuasion indicating that one possesses certain capabilities; and physiological states by which one judges capability, strength, and vulnerability. Although performance accomplishments are considered to be the most robust source of self-efficacy information, vicarious learning is also a powerful source (Bandura, 1986; 1997). That is, viewing others successfully accomplish a particular task can increase learners' perceptions of others' efficacy as well as their own efficacy for performing similar tasks (Bandura, 1997).

Vicarious learning experiences have been shown to enhance student-teachers' self-efficacy for using computers in their teaching. In 1993, Handler conducted a study with 133 education graduates. Participants responded to a survey regarding their perceptions of the value of pre-service computer experiences to their professional preparation. Results showed that observing cooperating teachers use computers during the student teaching experience was one of the three most important factors that influenced feelings of preparedness for the use of computers for instruction in their own classrooms. Downes (1993) investigated student teachers' uses of computers during practicum sessions in order to identify relationships among computer uses and specific practicum factors. Results indicated a significant increase in computer use over the three practicum sessions and this increase was consistent from one practicum group to the next. What is interesting to note, however, is that when examining the factors involved in the practicum environment, only one factor, supervising teachers' uses of computers with children, was significant. No other practicum-related factors, including level taught and technology resources, were significant. The influence of the supervising teachers' uses of computers was so strong that first-year students, whose supervising teachers used computers with children, were more likely to use computers with children than third-year students whose supervising teachers did not. Apparently, observing positive role models (in this case, supervising teachers) favorably influenced the student teachers to perform similarly.

While novice learners can acquire skills and strategies from social modeling, when performing independently they are likely to over- or underestimate their own capabilities (Schunk, 2001). However, students' judgments of progress, as well as their judgments of self-efficacy, increase in both accuracy and strength when goals are made explicit (Schunk, 2001). Research in education (Schunk, 1990) and organizational management (Lee, Locke, & Latham, 1989) has emphasized the importance of goals directed toward specific performance levels, with concrete, measurable outcomes. Because specific goals help define what constitutes an acceptable level of performance, the explication of these goals can help students make more accurate, as well as more robust, judgments of efficacy (Schunk, 2001). By establishing goals, students typically experience a sense of efficacy for attaining them.

The literature has established independent effects of both vicarious learning experiences and goal setting on learners' judgments of self-efficacy, yet little work has been done to examine how these strategies might be combined to create even more accurate and more robust judgments of efficacy. In 1992, Gist and Mitchell identified three general strategies for enhancing self-efficacy beliefs. Of these three, two related to vicarious learning and goal setting, respectively: providing opportunities to observe experts' practice and providing opportunities to address a specific goal while resolving a particular teaching issue. Gist and Mitchell concluded that these strategies contributed to building teachers' confidence for achieving effective teaching.

According to Neck and Manz (1992), when individuals mentally rehearse a task, they see themselves performing it and thus are exposed to the positive effect of modeling (i.e., learn through vicarious experiences). Furthermore, the intense cognitive processing that occurs during mental practice can heighten awareness of how to attain specific goals and hence increase goal commitment and task performance. Based on these premises, it was hypothesized that vicarious learning experiences and goal setting could be combined to achieve a significant effect on learners' self-efficacy beliefs and task performance.

Purpose of the Study

This study was designed to examine the impact of vicarious learning experiences and goal setting on preservice teachers' self-efficacy for technology integration. Specifically, this study was guided by the following research question:

What are the effects of vicarious experiences and goal setting on preservice teachers' judgments of self- efficacy for technology integration?

Based on the self-efficacy literature described above, it was hypothesized that preservice teachers who were exposed to vicarious experiences, related to successful technology integration, would experience significantly greater increases in judgments of self-efficacy for technology integration than those who were not exposed to these vicarious experiences. Furthermore, it was hypothesized that preservice teachers who were assigned specific goals would experience significantly greater increases in judgments of self-efficacy than those who were not assigned any goals. Finally, it was hypothesized that preservice teachers who were exposed to vicarious experiences *and* assigned specific goals would demonstrate the greatest increases in judgments of self-efficacy compared to students who received either one of these conditions alone.

Methods

Research Design

A 2 x 2 (Vicarious Experiences x Goal Setting) mixed factorial research design was used to examine how vicarious experiences and goal setting impacted preservice teachers' judgments of self-efficacy for technology integration. These independent variables were combined to form four experimental conditions: (a) NVE/NGS: no vicarious experiences and no goal setting (also defined as the control group), (b) NVE/GS: no vicarious experiences but with goal setting, (c) VE/NGS: vicarious experiences with no goal setting, and (d) VE/GS: vicarious learning experiences with goal setting.

Participants

Participation was solicited from the 408 students enrolled in an *Introduction to Education Technology* course during the spring of 2003. Among these students, 337 agreed to participate in the study, although complete data sets were available from only 280 participants, including 92 males and 188 females. The participants' ages ranged from 18 to 38 years (M = 19.88, SD = 2.69). The majority of the participants were freshmen (n = 153); the rest were sophomores (n = 72), juniors (n = 36), seniors (n = 16), and graduate students (n = 3). Participants were majoring in elementary education (n = 105), secondary education (n = 113) within various content areas, pre-kindergarten to kindergarten education (n = 13), and others (n = 49). The demographic data collected from the participants also showed that among the 280 participants, 268 students (96%) planned to become teachers after graduating. The majority (n = 221) of the participants had never taken a computer class before. Those participants who had completed previous computer classes reported that the classes were mostly introductory computer literacy courses. Thus, most participants in this study did not possess high levels of computer skills. Based on a 4-point Likert-style question (1-not confident, 4-very confident) participants' initial confidence levels averaged 2.77 with a standard deviation of 0.82. In general, the participants rated themselves somewhat confident to confident in their ability to use technology to teach.

The participants' pre-course understandings of computer uses and technology integration in teaching were illustrated by their responses to questions about "specific strategies to develop knowledge and skills for teaching with technology" and "important things to consider when planning technology use." Almost every participant emphasized the role of taking a lot of computer classes in order to improve their computer skills. Participants believed that abundant practice with computers and familiarization with computer programs would sufficiently prepare them for teaching with technology. Only about 20 participants (less than 10%) mentioned the role of cooperating with other teachers, and even fewer (n = 4) mentioned things such as experimenting with different software programs in order to see which worked best for students' learning and to integrate technology into lesson plans where relevant. Based on participants' pre-course responses, little understanding of computer uses and technology integration in teaching was apparent.

Instrument

A Likert-style survey measuring participants' self-efficacy beliefs for technology integration was developed by the first author and used as pre- and post-survey measures. The survey (see Appendix A) included 21 items regarding participants' confidence for technology use. The participants were asked to rate their levels of agreement (from 1-strongly disagree to 5-strongly agree) with statements related to their possession of confidence regarding technology use (e.g., "I feel confident that I understand computer capabilities well enough to maximize them in my classroom." "I feel confident I can regularly incorporate technology into my lessons, when appropriate to student learning.").

Content and construct validity of the instrument were examined. The evidence of content validity is primarily judgmental in nature and therefore was gathered prior to the actual administration of the instrument. In the process of instrument development, the first author established a panel consisting of six content experts in the area of self-efficacy (five professors and one graduate student). The experts were provided with a bibliography and a summary of the literature review. These served as the content universe. Individually, the experts reviewed the materials and commented on the adequacy of the conceptual definition. The first author also developed a rating sheet so that the experts could rate and make suggestions for each item on the instrument. With the feedback obtained from the experts' ratings, appropriate revisions of the instrument were made. It was believed that the content validity of the instrument was convincing.

The evidence of construct validity is primarily empirical in nature. In this study, evidence was gathered after the self-efficacy survey had been administered to the participants. A factor analysis was conducted on both pre-survey data and post-survey date to identify, from the 21 items, subsets of those items that could be clustered together to form constructs (i.e., factors). The factor analysis was exploratory in nature and was conducted to determine if the instrument developed for this study actually measured some meaningful constructs. As suggested by Gable and Wolf (1993), the number of factors and the relationship of items to factors were determined by the analysis rather than by the instrument developer's theoretical predictions.

The factor analysis of the pre-survey data produced a two-factor solution that explained 55.36% of the systematic covariance among the items (see Appendix B). The first factor (eigenvalue = 9.85) accounted for 46.92% of the covariance and consisted of sixteen items with loadings ranging from .51 to .84. The items defining this factor represented computer technology capabilities and strategies. The second factor (eigenvalue = 1.77) accounted for 8.4% of the covariance and consisted of five items with loadings ranging from .56 to .77. The five items represented external influences of computer technology uses (e.g., student assessment, restraints, oppositions, etc.). Therefore, it was determined that this second factor was not closely related to computer technology integration. Thus, the final version of the instrument only consisted of the sixteen items in Factor One. Another factor analysis was then conducted with the sixteen items on the post-survey data. A one-factor solution was suggested (see Appendix C). This factor (eigenvalue = 9.58) explained a total of 59.86% of the systematic covariance. Therefore, it was further confirmed that these sixteen items formed a valid instrument measuring a single construct.

Cronbach alpha coefficients were calculated for both pre-survey data and post-survey data (on the sixteen items) to determine the reliability of the instrument. Alpha coefficients of .94 (for pre-survey) and .96 (for post-survey) indicated that the instrument was highly reliable. Thus, the obtained factor solution and resulting reliability coefficients for the self-efficacy for technology integration scale suggests that the instrument exhibits construct validity and reliability. As such, it may be concluded that the resulting form of the instrument holds promise for its use in further research.

Procedures

The 18 lab sections of the course were randomly assigned to the four experimental conditions. During the third week of the semester, participants completed the demographic questionnaire which included 17 questions asking for information about age, gender, major, previous computer classes, self-judgment of confidence for using computers in teaching, and understandings of computer uses and technology integration in teaching. Some questions were in the format of multiple-choice, while others were short-answer, open-ended questions. The self-efficacy survey was also administered as a pre-course measure. Both the demographic questionnaire and the pre-survey were administered online, using the existing course management system, WebCT.

During the sixth week of the semester, the researcher (the first author) described the study to the students and solicited participation. Students who agreed to participate in the study signed the informed consent form and spent a regular two-hour lab session completing the assignments for this study, under the direction of the researcher. The participants were offered five points toward their course grades for participating. During the experiment, the participants worked actively, following closely the directions they received from the researcher. Almost all participants spent at least one hour viewing either the VisionQuest© CD-ROM or the WebQuest website to which they were assigned. During this time, they made notes or wrote responses on the worksheets provided. After the experiment, some participants spoke with the researcher expressing their great interest in the CD-ROM and trying to obtain more information about the content of the software as well as the process by which the software was developed. At the end of the experiment, the participants completed the self-efficacy survey for technology integration as a post-measure, once again accessing the survey online via the course WebCT.

In this study, vicarious experiences for technology integration were presented to the students using VisionQuest©, an instructional CD-ROM that features the technology practices and beliefs of six K-12 teachers. According to Ertmer, Conklin, Lewandowski, Osika, Selo, and Wignall (2003), "VisionQuest© is designed provide opportunities for users to explore models of effective technology integration" (p. 100). The various cases highlighted on the CD-ROM illustrate that technology integration can be achieved successfully in a variety of contexts despite differences in settings, resources, and student backgrounds.

VisionQuest© provides vicarious learning experiences for the user through the use of video segments, augmented by electronic artifacts (lesson plans, student products) from teachers' classrooms.

Cases are constructed such that users can explore teachers' classrooms either one at a time (case by case) or thematically (i.e., comparing components of technology integration across cases). Each case contains a variety of elements that combine to illustrate how teachers' visions for technology use are translated into practice. Users examine how teachers planned for integration, how they currently implement technology within their classrooms, and how they assess the impact of their efforts.

For the groups that engaged in vicarious experiences, that is, the VE/GS and VE/NGS groups, participants explored the VisionQuest© CD-ROM and observed the technology uses and classroom management strategies of the featured teachers. For the groups that engaged in goal setting, that is, the VE/GS and NVE/GS groups, participants were given a number of specific goals. For example, the following goals were assigned to the participants in the VE/GS condition:

> While you are exploring VisionQuest©, it helps to keep in mind what you are trying to do. A list of expected outcomes from this activity is shown on this page and can be thought of as goals that you are trying to accomplish. So while you are going through the VisionQuest©, you should keep in mind the following goals:
> For each teacher on VisionQuest©, determine:
> - his/her beliefs about technology use
> - the roles technology plays
> - the way he/she organizes technology-based class activities
> - the way students are assessed

Students in the NVE/GS condition, who viewed a variety of WebQuests instead of exploring VisionQuest©, received the following goals:

> While you are exploring these WebQuests, it helps to keep in mind what you are trying to do. A list of expected outcomes from this activity is shown on this page and can be thought of as goals that you are trying to accomplish. So while you are going through the WebQuests, you should keep in mind the following goals:
> For each WebQuest that you view, determine:
> - the instructional goal(s)
> - the procedure of how the students might achieve the goal(s)
> - how students' achievement of the goal(s) will be evaluated

For the groups that did not engage in vicarious learning experiences, that is, the NVE/GS and NVE/NGS groups, participants explored a website containing links to various WebQuests selected for this study. The WebQuests had the content of technology in teaching but did not contain characteristics of vicarious learning. For the groups that did not engage in goal setting, that is, the VE/NGS and NVE/NGS groups, participants received only instructions on how to navigate the VisionQuest© software or the WebQuest website, but nothing related to what knowledge/information they were expected to gain from the software or the website (i.e., no specific learning goals were assigned to these participants). In addition, in order to make sure that the participants would attend to the task they were required to accomplish, that is, exploring the information in the software or on the website rather than wasting time on irrelevant tasks, each participant was required to complete a worksheet related to the learning goals or the content of their specific tasks, depending on the experimental conditions in which they participated.

Data Analysis Strategies

Demographic data were analyzed in order to describe the characteristics of the participants in this study. Ranges, frequencies, means, and standard deviations were calculated for descriptive data. Responses to opened-ended questions were analyzed using the standard procedures of pattern seeking that are most commonly adopted by qualitative researchers.

Means and standard deviations were calculated for each experimental group on the pre- and post-surveys of self-efficacy for technology integration. ANOVA results [F (3, 276) = 0.06, p = .9818] showed that there were no significant differences among the means of the pre-survey scores for the four experimental groups. In order to see the effects of vicarious learning experiences and goal setting on preservice teachers' self-efficacy for technology integration, two-way ANOVA was used to analyze post-survey data. Furthermore, the ANOVA was also used to determine whether there were statistically significant differences among the means of the four experimental groups.

Results

Table 1 presents the mean scores and the standard deviations on the pre- and post-surveys of the four experimental groups. As noted above, differences among the means and standard deviations from the pre-survey scores were not significant. For the post-survey data, the VE/GS group, vicarious experiences with goal setting, had the highest mean score on ratings of self-efficacy for technology integration. The NVE/NGS control group, no vicarious experiences with no goal setting, had the lowest mean score. The control group had the largest standard deviation while the VE/GS group had the smallest standard deviation.

Table 1. *Means and Standard Deviations for the Pre- and Post-survey Scores on Self-Efficacy for Technology Integration*

Experimental Group	N	Mean		Standard Deviation	
		Pre-survey	Post-survey	Pre-survey	Post-survey
VE-GS	71	3.79	4.35	0.58	0.46
VE-NGS	70	3.82	4.06	0.62	0.52
NVE-GS	71	3.82	4.00	0.55	0.47
NVE-NGS	68	3.82	3.79	0.54	0.59

Table 2 presents the results of the two-way ANOVA for the post-survey scores on self-efficacy for technology integration. Significant main effects suggest that vicarious learning experiences [F (1, 276) = 25.63, $p < .0001$] and goal setting [F (1, 276) = 16.59, $p < .0001$)] significantly increased participants' self-efficacy for technology integration. No significant interaction [F (1, 276) = 0.38, $p = .5388$] between the two independent variables was found, suggesting that the effect of vicarious learning experiences on self-efficacy beliefs did not change at the different levels of goal setting, and vice versa. Thus, we can conclude that vicarious learning experiences and goal setting both increased the participants' self-efficacy, but when vicarious learning experiences and goal setting were combined, the increase was the greatest. In addition, the effect size index f was calculated following the method presented by Cohen (1988). It was found that, with $f = .40$, the effect size was large according to Cohen's definition (1988, p. 284-288), meaning that the systematic variances were largely explained by the specific conditions to which the participants in this study were assigned to.

Table 2. *Two-way ANOVA for the Post-survey Scores on Self-Efficacy for Technology Integration*

Source	*df*	Type III SS	Mean Square	*F*	*p*	ES Index *f*
VE-NVE	1	6.7601	6.7601	25.63	< .0001	.40
GS-NGS	1	4.3770	4.3770	16.59	< .0001	
VE-NVE*GS-NGS	1	0.0999	0.0999	0.38	.5388	

The ANOVA results also demonstrate the pair-wise differences among the four experimental groups: The mean score of the group in which both vicarious learning experiences and goal setting were present was significantly higher than the mean score of the control group ($p < .0001$). In addition, the mean score of the group in which both vicarious learning experiences and goal setting were present was significantly higher than the mean scores of the groups where only vicarious learning experiences ($p = .0010$) or goal setting ($p < .0001$) was present. The mean score of the group where only vicarious learning experiences was present was significantly higher than the mean score of the control group ($p = .0020$), and the mean score of the group where only goal setting was present was also significantly higher than the mean score of the control group ($p = .0155$). Figure 1 visually presents the mean score differences.

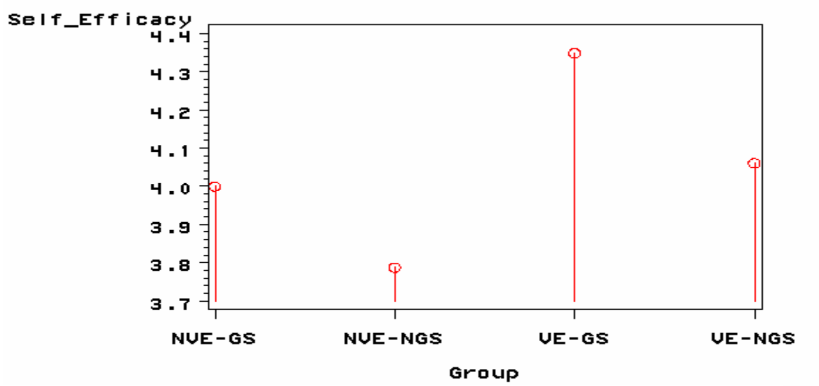


Figure 1. Mean score differences of the four experimental groups for the post-survey

Discussion

The results of this study indicated that preservice teachers who were exposed to vicarious experiences that were related to successful technology integration (with and without goal setting) experienced significantly greater increases in judgments of self-efficacy for technology integration than those who were not exposed to these vicarious experiences. These results support previous research regarding the benefits of vicarious learning on judgments of self-efficacy (Albion, 1996; Ertmer et al., 2003; Downes, 1993; Handler, 1993) and highlight the potential benefit to providing preservice teachers with opportunities for observing exemplary technology-using teachers as one way to increase their self-efficacy for effectively using technology in their own classrooms. Specifically, in this study, vicarious learning experiences were provided via an electronic instructional tool, VisionQuest© CD-ROM. This type of modeling helps eliminate logistic problems that might be associated with direct classroom observations and can be easily incorporated into a teacher education program as either a self-paced reflection tool or an instructor-led class activity. Other forms of electronic vicarious learning experiences, such as those provided by videos and Internet webpages, may also bring about similar effects on self-efficacy beliefs for technology integration.

The results of this study also indicated that preservice teachers who used specific goals, with and without vicarious experiences, experienced significantly greater increases in judgments of self-efficacy for technology integration than those who were not assigned any goals. These results also support previous research regarding the benefits of goal setting on self-efficacy beliefs (Bandura, 1997; Schunk, 2001; Schunk & Ertmer, 1999). Thus, providing preservice teachers with goals seems to be an effective way to enhance efficacy levels for technology integration, which can be easily achieved by making the link between class objectives and learning goals explicit.

More importantly, the results of this study showed that preservice teachers who were exposed to vicarious learning experiences and who were assigned specific goals experienced significantly greater increases in judgments of computer self-efficacy than those who received only one of these two conditions. These results support the suggestion made by others (Gist & Mitchell, 1992; Neck & Manz, 1992) regarding effective strategies for increasing self-efficacy as well as the possible benefit to be gained by combining strategies. As such, teacher educators might consider using both strategies when helping preservice teachers learn about technology integration. For example, instructors might anticipate increases in students' self-efficacy for technology integration when exemplary uses of technology in K-12 classrooms are presented and students explore these uses according to specific goals. What makes this practice more desirable is that both the exemplary uses and goals can be easily incorporated into software packages or other forms of electronic delivery.

Limitations and Suggestions for Future Work

The characteristics of the participants in this study may limit generalizability to participants with different characteristics. For example, the participants in this study were primarily female residential undergraduate students at the beginning of their teacher education programs (i.e., 67% female with an average age of less than 20, with over half being freshmen). In addition, participants did not demonstrate much initial understanding of computer uses and technology integration in teaching. These characteristics of the participants would make it difficult to generalize the results of this study to preservice teachers in other programs who might have different characteristics, although by carefully describing the characteristics of the participants would help make the generalization possible.

Another unique feature of this study was that the experiment was administered to the participants during a regular two-hour lab session of the course, which would be considered to be a relatively short treatment time. Therefore, there would be no indication of long-term, lasting effects of the treatment from the results of this study.

In addition, the format of the goals used in this study lent themselves to advanced organizer. Therefore, it is arguable that the effect observed on self-efficacy beliefs might be, at least in part, due to the role of cognitive scaffolding these goals played in affecting participants' self-efficacy. It should be advisable that the cognitive scaffolding features of these goals be removed in order to see if the findings of this study are replicable.

Therefore, for future research work, considerations might be made to administer the experiment over a longer period of time, say for a whole semester, in order to investigate the long-term effects of vicarious learning experiences and goal setting. If any differences or interesting patterns were found in preservice teachers' self-efficacy beliefs compared with the one-time brief exposure to vicarious learning experiences and goal setting, meaningful implications might be made to the benefit of teacher education programs. The results of this study should also be further qualified by trying out the same research design and experiment on preservice teachers in different types of teacher education programs where students' characteristics vary from one to another. Lastly, it would be important to examine the actual technology use by the students who have achieved high levels of computer self-efficacy, to verify the assumed relationship between high efficacy and actual classroom use.

Conclusions

The findings of this study contribute to the existing body of literature in two significant ways: (1) this study confirmed the findings from previous research that demonstrated how preservice teachers benefit from observing teacher models presented via vicarious learning experiences, such as those provided by VisionQuest© (Ertmer et al., 2003), and from setting learning goals for mastering computer-related tasks (Schunk & Ertmer, 1999), and (2) this study provided support for the hypothesized combined effect of vicarious learning experiences and goal setting on increasing preservice teachers' self-efficacy for technology integration. Though enhanced self-efficacy beliefs do not automatically translate into the actual use of technology among teachers, they are a necessary condition for technology integration. Teachers' self-efficacy beliefs have been found to be useful indicators of likely success at technology integration (Olivier & Shapiro, 1993). Furthermore, increased performance with computer-related teaching practices has also been found to be significantly related to higher levels of computer self-efficacy (Harrison, Rainer, Hochwarter, & Thompson, 1997). Therefore, from the perspective of teacher educators, the use of electronic vicarious learning experiences and the incorporation of learning goals may help preservice teachers develop the confidence they need to become effective technology users within their own classrooms. Thus, as our future teachers achieve high confidence levels for technology integration and develop powerful strategies for technology implementation, meaningful technology use can come closer to being the norm, rather than the exception, in our K-12 classrooms.

References

Albion, P. R. (1996). Student teachers' use of computers during teaching practice in primary classrooms. *Asia-Pacific Journal of Teacher Education, 24*, 63-73.

Albion, P. R. (1999). *Self-efficacy beliefs as an indicator of teachers' preparedness for teaching with technology*. Retrieved July 25, 2002, from http://www.usq.edu.au/users/albion/papers/site99/1345.html

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Becker, H. J. (2000). Access to classroom computers. *Communications of the ACM. 43*(6), 24-25.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Compeau, D., Higgins, C. A., & Huff, S. (1999). Social cognitive theory and individual reactions to computing technology: A longitudinal study. *MIS Quarterly, 23*, 145-158.

Downes, T. (1993). Student-teachers' experiences in using computers during teaching practice. *Journal of Computer Assisted Learning, 9*, 17-33.

Eachus, P., & Cassidy, S. (1999). *Developing the computer self-efficacy (CSE) scale: Investigating the relationship between CSE, gender and experience with computers*. University of Salford, United Kingdom, 1-10. Retrieved November 12, 2002, from http://www.salford.ac.uk/healthSci/selfeff/selfeff.htm

Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development, 47*(4), 47-61.

Ertmer, P. A., Conklin, D., Lewandowski, J., Osika, E., Selo, M., & Wignall, E. (2003). Increasing preservice teachers' capacity for technology integration through use of electronic models. *Teacher Education Quarterly, 30*(1), 95-112.

Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain: measuring attitudes and values in corporate and school settings* (2nd ed.). Norwell, MA: Kluwer Academic.

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*, 183-211.

Handler, M. G. (1993). Preparing new teachers to use computer technology: Perceptions and suggestions for teacher educators. *Computers and Education, 20*, 147-156.

Harrison, A. W., Rainer, R. K., Jr., Hochwarter, W. A., & Thompson, K. R. (1997). Testing the self-efficacy-performance linkage of social-cognitive theory. *Journal of Social Psychology, 137*, 79-87.

Lee, Y. E., Locke, E. A., & Latham, G. P. (1989). Goal setting theory and job performance. In L. A. Pervin (Eds.), *Goal concepts in personality and social psychology* (pp. 291-326). Hillsdale, NJ: Erlbaum.

Marcinkiewicz, H. R. (1996). Motivation and teachers' computer use. (Eric Document Reproduction Service No. ED 397 818)

Neck, C. P., & Manz, C. C. (1992). Thought self-leadership: The influence of self-talk and mental imagery on performance. *Journal of Organizational Behavior, 13*, 681-699.

Oliver, T. A., & Shapiro, F. (1993) Self-efficacy and computers. *Journal of Computer-Based Instruction. 20*, 81-85.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, *25*, 71-86.

Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives*, 2nd ed. (pp. 125-151). Mahwah, NJ: Lawrence Erlbaum.

Schunk, D. H., & Ertmer, P. A. (1999). Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology, 91*, 251-260.

Zehr, M. A. (1997). Teaching the teachers. *Education Week: Technology Counts, 17*(11), 26-29.

Zehr, M. A. (1998). The state of the states: Many still haven't dealt with the most difficult policy issues. *Education Week: Technology Counts, 18*(5), 69-71.

Appendix A
Computer Technology Integration Survey

Direction:

The purpose of this survey is to determine how you feel about integrating technology into classroom teaching. For each statement below, indicate the strength of your agreement or disagreement by circling one of the five scales.

Below is a definition of technology integration with accompanying examples:

<u>Technology integration</u>:

Using <u>computers</u> to support students as they construct their own knowledge through the completion of authentic, meaningful tasks.

<u>Examples</u>:

- Students working on research projects, obtaining information from the Internet.
- Students constructing Web pages to show their projects to others.
- Students using application software to create student products (such as composing music, developing PowerPoint presentations, developing HyperStudio stacks).

Using the above as a baseline, please circle one response for each of the statements in the table:

SD = Strongly Disagree, D = Disagree, NA/ND = Neither Agree nor Disagree, A = Agree, SA = Strongly Agree

1.	I feel confident that I understand computer capabilities well enough to maximize them in my classroom.	SD	**D**	NA/ND	A	SA
2.	I feel confident that I have the skills necessary to use the computer for instruction.	SD	**D**	NA/ND	A	SA
3.	I feel confident that I can successfully teach relevant subject content with appropriate use of technology.	SD	**D**	NA/ND	A	SA
4.	I feel confident in my ability to evaluate software for teaching and learning.	SD	**D**	NA/ND	A	SA
5.	I feel confident that I can use correct computer terminology when directing students' computer use.	SD	**D**	NA/ND	A	SA
6.	I feel confident I can help students when they have difficulty with the computer.	SD	**D**	NA/ND	A	SA
7.	I feel confident I can effectively monitor students' computer use for project development in my classroom.	SD	**D**	NA/ND	A	SA
8.	I feel confident that I can motivate my students to participate in technology-based projects.	SD	**D**	NA/ND	A	SA
9.	I feel confident I can mentor students in appropriate uses of technology.	SD	**D**	NA/ND	A	SA
10.	I feel confident I can consistently use educational technology in effective ways.	SD	**D**	NA/ND	A	SA
11.	I feel confident I can provide individual feedback to students during technology use.	SD	**D**	NA/ND	A	SA

12.	I feel confident I can regularly incorporate technology into my lessons, when appropriate to student learning.	SD	**D**	NA/ND	A	SA
13.	I feel confident about selecting appropriate technology for instruction based on curriculum standards.	SD	**D**	NA/ND	A	SA
14.	I feel confident about assigning and grading technology-based projects.	SD	**D**	NA/ND	A	SA
15.	I feel confident about keeping curricular goals and technology uses in mind when selecting an ideal way to assess student learning.	SD	**D**	NA/ND	A	SA
16.	I feel confident about using technology resources (such as spreadsheets, electronic portfolios, etc.) to collect and analyze data from student tests and products to improve instructional practices.	SD	**D**	NA/ND	A	SA
17.	I feel confident that I will be comfortable using technology in my teaching.	SD	**D**	NA/ND	A	SA
18.	I feel confident I can be responsive to students' needs during computer use.	SD	**D**	NA/ND	A	SA
19.	I feel confident that, as time goes by, my ability to address my students' technology needs will continue to improve.	SD	**D**	NA/ND	A	SA
20.	I feel confident that I can develop creative ways to cope with system constraints (such as budget cuts on technology facilities) and continue to teach effectively with technology.	SD	**D**	NA/ND	A	SA
21.	I feel confident that I can carry out technology-based projects even when I am opposed by skeptical colleagues.	SD	**D**	NA/ND	A	SA

Appendix B

Factor Analysis Results and Factor Loadings: Pre-Survey Scores (N=280)

	Item Number	Stem	Loading
Factor 1 (alpha = .94)			
Computer Technology Capabilities and Strategies	1	I feel confident that I understand computer capabilities well enough to maximize them in my classroom.	.84
	6	I feel confident I can help students when they have difficulty with the computer.	.79
	2	I feel confident that I have the skills necessary to use the computer for instruction.	.79
	5	I feel confident that I can use correct computer terminology when directing students' computer use.	.75
	4	I feel confident in my ability to evaluate software for teaching and learning.	.72
	3	I feel confident that I can successfully teach relevant subject content with appropriate use of technology.	.71
	9	I feel confident I can mentor students in appropriate uses of technology.	.69
	7	I feel confident I can effectively monitor students' computer use for project development in my classroom.	.68
	11	I feel confident I can provide individual feedback to students during technology use.	.64
	10	I feel confident I can consistently use educational technology in effective ways.	.61
	18	I feel confident I can be responsive to students' needs during computer use.	.60
	14	I feel confident about assigning and grading technology-based projects.	.59
	12	I feel confident I can regularly incorporate technology into my lessons, when appropriate to student learning.	.56
	13	I feel confident about selecting appropriate technology for instruction based on curriculum standards.	.56
	16	I feel confident about using technology resources (such as spreadsheets, electronic portfolios, etc.) to collect and analyze data from student tests and products to improve instructional practices.	.55
	8	I feel confident that I can motivate my students to participate in technology-based projects.	.51
Factor 2 (alpha = .75)			
External Influences of Computer Technology Uses	21	I feel confident that I can carry out technology-based projects even when I am opposed by skeptical colleagues.	.77
	20	I feel confident that I can develop creative ways to cope with system constraints (such as budget cuts on technology facilities) and continue to teach effectively with technology.	.65
	15	I feel confident about keeping curricular goals and	.64

	technology uses in mind when selecting an ideal way to assess student learning.	
19	I feel confident that, as time goes by, my ability to address my students' technology needs will continue to improve.	.64
17	I feel confident that I will be comfortable using technology in my teaching.	.56

Appendix C

Factor Analysis Results and Factor Loadings: Post-Survey Scores (N=280)

	Item Number	Stem	Loading
Factor 1 (alpha = .96)			
Computer Technology Capabilities and Strategies	2	I feel confident that I have the skills necessary to use the computer for instruction.	.81
	3	I feel confident that I can successfully teach relevant subject content with appropriate use of technology.	.80
	14	I feel confident about assigning and grading technology-based projects.	.80
	10	I feel confident I can consistently use educational technology in effective ways.	.79
	8	I feel confident that I can motivate my students to participate in technology-based projects.	.79
	6	I feel confident I can help students when they have difficulty with the computer.	.79
	9	I feel confident I can mentor students in appropriate uses of technology.	.78
	11	I feel confident I can provide individual feedback to students during technology use.	.78
	18	I feel confident I can be responsive to students' needs during computer use.	.78
	12	I feel confident I can regularly incorporate technology into my lessons, when appropriate to student learning.	.78
	1	I feel confident that I understand computer capabilities well enough to maximize them in my classroom.	.77
	16	I feel confident about using technology resources (such as spreadsheets, electronic portfolios, etc.) to collect and analyze data from student tests and products to improve instructional practices.	.75
	13	I feel confident about selecting appropriate technology for instruction based on curriculum standards.	.74
	4	I feel confident in my ability to evaluate software for teaching and learning.	.74
	5	I feel confident that I can use correct computer terminology when directing students' computer use.	.73
	7	I feel confident I can effectively monitor students' computer use for project development in my classroom.	.73

Investigate the Impact of a Web-Based Learning Environment on Student Motivation and Achievement of Learning Science Concepts

Shiang-Kwei Wang
New York Institute of Technology

Chia-chi Yang
University of Missouri, Columbia

Abstract

A high school science teacher reported that the students have motivation and learning problems to understand the concept of fossilization. Working with the science teacher over eighteen months, Fossilization Web-Based Learning Environment (Web-LE, http://iris.nyit.edu/~skwang/fossil) was designed and developed by a group of students in the Department of Instructional Technology at University of Georgia to investigate the impact of the this tool on student motivation and learning achievement. The results showed that this Web-LE provided an appropriate tool for the teacher to teach the unit of fossilization, enabled the teacher to design specific assessment strategies, and improved student motivation. The teacher will use the tool in the future class since Fossilization Web-LE is customized for him and it is proved useful in his situation-specific context.

Introduction

The primary setting of this developmental research was a local private school located in a small city in northeast Georgia. This school is equipped with well-organized computer technology including computer labs, laptop, and wireless network. Students and teachers are used to the computer environment and equipped with relatively higher computer literacy. Some teachers in this school looked for strategies of integrating technology into their classroom. One tenth grade science teacher reported that he needed assistance to develop interactive instructional materials in order to solve his instructional problems. The team from the Department of Instructional Technology at University of Georgia analyzed his problems and decided to work with him. Twenty two months were invested on this project to conduct a developmental research and design Fossilization Web-LE. Fossilization Web-LE was designed to utilize the unique features and elements of the World Wide Web (WWW) to enhance and sustain the motivation of learners in the context of secondary science education. The project falls within the category of developmental research goals described by Reeves (2000). The purpose of developmental research is twofold: to focus on developing approaches to improve situation-specific teaching and learning problems and to generate methodological directions for design and evaluation for future development and implementation efforts.

Instructional Problems

The science teacher reported that the students have problems to understand the concept of fossil formation. Fossilization is a result of complex combination of an organism, ecological condition and physical burial. The fossil would be formed only under the correct situation. Students have to understand that how and why ecological conditions and physical burials influence fossilization. The teacher couldn't find accurate or useful materials to help students visualize the process of fossilization. Before the development of this tool, the teacher demonstrated volcanic films or pictures to help students to have a big picture of the fossilization conditions. He also asked students to pick up one condition and represent it by drawing. However, students could not understand the overall difference among the potential conditions of developing fossils. He needed moving illustrations describing different conditions of fossilization but he could not find appropriate materials to do so. The other problem was that students had relatively low motivation when learning the unit of fossilization. The science teacher needed a tool that is realistic and accurate to describe the process of fossil formation and provide opportunities for learners to consider the potential combinations of different decision. The learner will be able to identify situations that will cause fossilization.

Instructional Design Strategies

Several strategies for increasing students' intrinsic motivation will be involved in the Web-LE. First, Lepper and Hodell's (1989) four characteristics (challenge, curiosity, control, and fantasy) of tasks that promote individual intrinsic motivation were integrated into the instructional design. Second, raises and sustains learners' motivation by integrating multimedia objects into the learning context. Learning with multimedia provides an effective alternative instructional strategy (Mathewson, 1999). From the motivational perspective, using medium to assist instruction indeed enhance learners' interest (Freeley, 1982; Kramarski & Feldman, 2000). The multimedia objects were employed to represent the fossil concepts to enhance intrinsic motivation. Third, the teacher designed a task requiring higher order thinking skills for the unit of fossilization and encourages students to explore all scenarios to generate the conclusions instead of learning by rote.

Overviews of the System

After completing learning the concept of fossilization with this cognitive tool, students will be able to identify conditions necessary for fossilization and construct possible scenarios for fossil formation by manipulating the variables in the simulated processes. Fossilization Web-LE includes eighteen different scenarios to describe various conditions of fossilization and provides students the best way to observe most potential conditions to develop fossils. This is a three-day learning session. The teacher distributed the assignment in the first day and students had to solve problems within three days by exploring the scenarios provided the tool. These problems required critical thinking abilities such as comparing the similarities and differences and infer the conclusion.

When entering the Web-LE, learners will see the instructional page to explain the learning objectives and goals of the fossil unit as well as how to use the tool under the "manual" section .Figure 1 is the flow chart of fossilization cognitive tool created in Flash. After reading the task of exploring this interactive learning model, students will be prompted by the main interface of this tool. They have to follow the order of selecting an organism (first), an ecological condition (second), and then a physical burial (third). The combination of three decisions will decide whether the fossil will be developed and the simulation of fossilization begins after the combination has been made. The dinosaur is the only available organism in this version. Some abstruse terms regarding fossilization were identified by the teacher and illustrated by animations and descriptions in the online glossary. Students can assess to the online glossary to look up detailed information if they have any questions about the scientific terms.

The system will record paths that learners just made and help learners to identify the learning progress.

The complete eighteen possible combinations including:

1. Organism: dinosaur
2. Ecological status (3 conditions): temperate rainforest, tropical rainforest and tropical mountains
3. Physical burial (6 conditions): weathering, ashfall, lava flow, pyroclastic, swamp, and flood

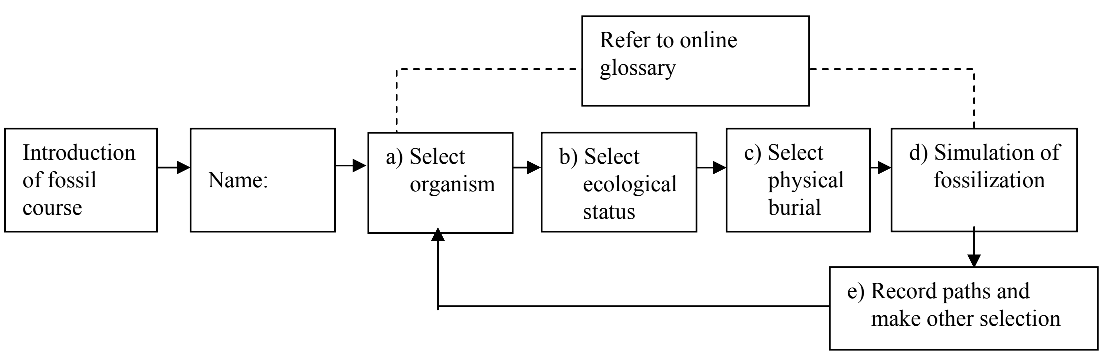


Figure 1. Flow chart of fossilization cognitive tool

Media development

The primary audience is high school students who are appeal to realistic graphics and animations, and their learning style is more visual oriented. The science teacher also requested for the realistic representation of the fossilization processes. Several endeavors were made to achieve the requirements. Referring to the accurate images, the dinosaur was developed with an accurate proportion and color. 3D software (Corel Bryce) was used to develop the realistic landscapes and objects including ancient plants, volcanoes, lakes and fossils in the fossilization animations. Flash MX (Macromedia) was the primary authoring tool because of its ability of integrating multimedia and the capability of optimizing media to enhance the speed of transmission on the Internet. The team met the teacher regularly and worked closely with him to conduct formative evaluation to ensure the graphical objects, descriptions of scenarios, information organized were accurate and realistic to the target learners.

Results

The teacher and twenty-seven tenth-grade students participated in this study, spending three class sessions using Fossilization Web-LE to finish cognitive tasks assigned by the teacher. Motivation questionnaire data were collected at the end of the third session allowing students to evaluate the tool and examine their attitudes toward using it. The student grades were also collected and analyzed to determine their learning performance. Observation, teacher interview, and twelve focus student interviews were collected during the sessions and after the three-day study was over.

The motivation questionnaire, observational protocol, interview questions were designed based on Newmann's identification (1992, p.13) of levels of engagement: (1) observable behavioral responses, (2) covert cognitive responses activated during learning, and (3) interest. The researcher carried out statistical analyses and correlation analyses to examine factors that might have influenced student motivation and learning performance. Newmann's levels of engagement helped the researcher scrutinize indicators of motivation that surfaced during the learning process. Student grades, student interviews, and the teacher interview were used to triangulate the data collected from the motivation questionnaire. The results showed that students reaped moderate learning and strong motivational benefit through using Fossilization Web-LE. Several minor interface design flaws were found during the implementation. These problems have since been fixed so that the teacher can use the revised program with future classes.

Lepper and Hodell's (1989) four factors for improving motivation were applied to the design of Fossilization Web-LE, and the tool had positive impact on student motivation as a result. Most students admitted that compared to other units, motivation toward learning fossilization was relatively lower without using the software; such statements reflected the teacher's statement of his instructional problems. Results collected from the motivation questionnaire and student interviews revealed an important message: student interest toward learning the topic of fossilization was higher when learning with Fossilization Web-LE. The results of implementation provided significant evidences that the tool solved the teacher problems. It enabled the teacher design various tasks that he could not do before using the tool and improved student motivation of learning the unit of fossilization.

Reference

Freeley, J. T. (1982). Content Interests and Media Preferences of Middle-Graders: Differences in a Decade. *Reading World*, 22(1), 11-16.

Lepper, M. R., & Hodell, M. (1989). Intrinsic motivation in the classroom. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 2, pp. 73-105). San Diego, CA: Academic Press.

Kramarski, B, & Feldman, Y. (2000). Internet in the classroom: Effects on reading comprehension, motivation and metacognitive awareness, *Educational Media International*, *37(3)*, 149-55.

Mathewson, J. (1999). Visual-spatial thinking: An aspect of science overlooked by educators, *Science Education*, *83(1)*, 33-54.

Newmann, F. M. (Ed.). (1992). *Student engagement and achievement in American secondary schools*. New York: Teachers College Press.

Reeves, T. C. (2000). Socially responsible educational technology research. *Educational Technology*, *40*(6), l9-28.

Acknowledgements

The fossilization Web-LE is developed by Shiang-Kwei Wang, Chia-chi Yang, Wan-Ting Huang, Beaumie Kim, and Dr. Hill under Dr. Reeves and Dr. Hill's supervise. We appreciate Jack Kridler's contribution to this project as a subject expert.

A Georgia Program for Preservice Teachers: Professional Development for Integrating TechnologyInto the K-12 Classroom

John R. Wiggins
Christa Harrelson
William Gillespie
University of Georgia

Paige Campbell
Georgia College and State University

Abstract

This study addresses how Georgia's InTech program served to train a cohort of pre-service teachers during their student teaching semester. The BellSouth Foundation provided the resources, through a grant, to study the impact of the InTech program, which offers extensive, curriculum-based professional training for incorporating technology into the Georgia K-12 curriculum. For this study, a sample of pre-service teachers were administered a pre- and post-training survey to indicate how InTech training affected the individuals. This survey especially focuses upon perceived individual technical skills on various items of technology and how comfortable the pre-service teachers felt in integrating technology in classroom curriculum. Using both quantitative and qualitative methods, the survey results indicate that pre-service teachers felt more comfortable in using technology overall and increased their perceived skills for particular items of technology. According to participant comments, InTech training provided these pre-service teachers with hands-on, relevant training that increased efficacy in incorporating technology into lessons.

InTech, a comprehensive, 50-hour training program, serves as the backbone of this project. InTech, Phase One of the Georgia Framework for INtegrating TECHnology, is designed to enhance the existing P-16 curriculum using modern technologies as a catalyst for fundamental changes in the teaching and learning process. The Georgia InTech Framework recognizes that instructional redesign is best accomplished through professional development activity that simultaneously builds teacher skills in five interrelated areas of proficiency. Georgia's Five Critical Areas for technology professional development target improved student achievement by: 1) focusing on Georgia's Quality Core Curriculum Standards, 2) using modern technological resources, 3) incorporating these technological resources into new designs for teaching and learning, 4) developing and using classroom management strategies which enable effective use of technology in the classroom, and, lastly, 5) blending these components into a new and enhanced classroom pedagogy.

Description of quantitative methods and findings - The questionnaire asks respondents to rate themselves on their ability to use equipment effectively in classroom instruction. Both the pre-test and post-test use a five-point scale for scoring answers from not comfortable at all to comfortable enough to use in classroom. The differences in responses for each question between the pre- and post-test are compared for statistical significance using a paired sample t-test. Quantitative findings give strong evidence that preservice teachers feel more technically competent on each piece of equipment after the training. Consistent with other studies on the topic, the largest gains reported by preservice teachers come from experiences less available in classrooms or home. The largest gains in proficiency occurred with digital cameras, capturing images from the Internet, and with scanners and opaque projectors. Other notable gains came from creating web-pages, using computer multimedia and zip drives, and using LCD projectors or videodisc players. Preservice teachers also registered an increase in the utilization of list-serves in the classroom. Smaller gains in student teacher proficiency included using overheads, video cameras, cd-roms, and 35mm cameras. These gains are positive and significant at the .05 level of confidence indicating that InTech training increased proficiency in these items as well.

Description of qualitative methods and findings - Participant responses to open-ended survey questions, journal entries, and e-mail responses were analyzed to provide a richer picture of the training.

Content analysis methods were used to identify the emerging themes from these sources. These themes included: the perceived relevancy of training, how it impacts hands on learning in the classroom, how teacher efficacy and growth is affected, and perceived changes in teacher development. These themes correlate to the quantitative findings. Data indicates a positive response to the InTech training. Teachers believe that they are better able to use technology to enhance both teaching and learning. Most preservice teachers desire experiences in their teacher preparation that they feel are timely and relevant to their professional development. Two areas of particular importance to the preservice teachers in this program are the marketability of their InTech training and the importance of technology in our society today. The most prevalent theme contained in the preservice teacher free responses about the training address their increased efficacy toward using technology in the classroom and their appreciation for the professional growth they experienced during the training. Over half of the preservice teachers expressed their increased efficacy toward integrating technology into the curriculum as a result of InTech training. They felt both more comfortable and more confident in their newly refined technology skills.

In conclusion, using both quantitative and qualitative methods, the survey results indicate that pre-service teachers felt more comfortable in using technology overall and increased their perceived skills for particular items of technology. According to participant comments, InTech training provided these pre-service teachers with hands-on, relevant training that increased efficacy in incorporating technology into lessons.

Introduction

This study addresses how Georgia's InTech program serves to train a cohort of pre-service teachers during their student teaching semester. The BellSouth Foundation provided the resources, through a grant, to study the impact of the InTech program, which offers extensive, curriculum-based professional training for incorporating technology into Georgia K-12 curriculum. For this study, a sample of pre-service teachers were administered a pre- and post-training survey to indicate how InTech training affected the individuals. This survey especially focuses upon perceived individual technical skills on various items of technology and how comfortable the pre-service teachers felt in integrating technology in classroom curriculum. Using both quantitative and qualitative methods, the survey results indicate that pre-service teachers felt more comfortable in using technology overall and increased their perceived skills for particular items of technology. According to participant comments, InTech training provided these pre-service teachers with hands-on, relevant training that increased efficacy in incorporating technology into lessons.

Study Rationale

The 1995 Needs Assessment for the State of Georgia indicates that over 40% of elementary, 37% of middle, and 25% of high school teachers give a "low" rating of their knowledge regarding the effective use of technology. In addition, approximately three-fourths of all elementary, middle, and high school teachers surveyed rate their level of access to technology-based inservice training as low or medium. Available training receives low marks according to these same teachers. Regarding the quality of technology related professional development activities, close to 80% of these teachers give a rating of low or medium. Over two-thirds of elementary, middle and high school teachers rate their administrators' knowledge of effective technology as low or medium. Thus, there seems to be a clear need for training that will provide teachers and administrators with the skills necessary to effectively integrate technology into the K-12 curriculum.

The results of the Georgia Needs Assessment regarding inservice training are consistent with the findings regarding preservice teacher education contained in the report titled *Technology and the New Professional Teacher: Preparing for the 21st Century Classroom*. This report, commissioned by NCATE (the National Council for the Accreditation of Teacher Education), found technology integration to be lacking or treated as an after-thought in most preservice teacher education programs. Brooks and Kopp, in their book titled *Handbook of Research on Teacher Education*, voiced similar concerns:

Undergraduate teacher-training institutions are not taking a convincing or focused leadership role in identifying solid evidence about the applications of technology to teacher training. The best and most consistent exposure for teachers to classroom-relevant technologies is often at the inservice or private sector level. In short, the information age has yet to significantly influence teacher training. (1990, p. 499)

As Brooks and Kopp state, most teachers graduate from preservice teacher education programs with little or no exposure to the effective integration of technology into the curriculum. They conclude that it is left to the teacher, once he or she reaches the classroom, to seek out technology training either through private training or through school-based inservice programs. And, unfortunately, as the Georgia Needs Assessment reveals, these inservice programs are not effectively providing the majority of the teachers in Georgia with the training that they need or desire. Schrum and Dehoney (1998), citing research by Siegel (1995) and others (Green & Gilbert, 1995), summarize the current status of preservice and inservice technology training. Preservice programs fall short, thus leaving newly hired teachers to seek out inservice training, yet studies find current inservice training to be unavailable or of poor quality when it is present. Schrum and Dehoney state:

In a recent study of technology staff development in schools (Siegel, 1995), it was reported that K-12 teachers are generally not receiving enough time, access, support, or encouragement to become comfortable with computers. Preservice education planners cannot expect teachers to gain the necessary skills on the job. (1998, p.24)

Smith (1999) offers the same assessment: "Most teachers are not adequately prepared to make use of exciting new educational technologies because neither their teacher education programs nor their schools have provided sufficient time or incentives for them to learn (p. 26)."

Professional Development Model

InTech, a comprehensive, 50-hour training program, serves as the backbone of this study. InTech, Phase One of the Georgia Framework for INtegrating TECHnology, is designed to enhance the existing P-16 curriculum using modern technologies as a catalyst for fundamental changes in the teaching and learning process. The Georgia InTech Framework recognizes that instructional redesign is best accomplished through professional development activity that simultaneously builds teacher skills in five interrelated areas of proficiency. Georgia's Five Critical Areas for technology professional development target improved student achievement by: 1) focusing on Georgia's Quality Core Curriculum Standards, 2) using modern technological resources, 3) incorporating these technological resources into new designs for teaching and learning, 4) developing and using classroom management strategies which enable effective use of technology in the classroom, and, lastly, 5) blending these components into a new and enhanced classroom pedagogy.

InTech goes beyond many technology-related staff development programs in both its intensity and duration, lasting approximately four months. This allows time for teachers to reflect on their experiences and further engage their students in technology connected activities. During this time, preservice teachers are collaborating with other cohort members, corresponding with their InTech instructors, planning with their cooperating teachers within their placement schools, and actively incorporating technology into the teaching and learning process.

InTech requires the cohort of preservice teachers to attend a day of training, develop and implement a technology connected lesson in the classroom, and then return to InTech training between two and four weeks later to share the experience. This continues throughout the 50-hour training program with each preservice teacher developing and implementing a minimum of four technology connected lessons. The seventh and final day of training culminates with a presentation by groups of preservice teachers highlighting and sharing what they have developed and implemented.

Data Collection and Analysis

Qualitative and quantitative methods are employed to provide depth to the description of what InTech offers to preservice students. The source for quantitative data was a questionnaire administered both pre and post training. Qualitative data were gathered from the questionnaire, journal responses, lesson plans and email correspondence.

Qualitative Methodology

To provide a better description of what InTech provides, participants provided responses through open-ended survey questions, journal entries, and e-mail responses. This data source provided a deeper understanding of what the students were experiencing as they endeavored to develop their technology connected lesson plans. Qualitative content analysis is used to gain several overall themes from these sources. These themes included the perceived relevancy of InTech training, how InTech impacts hands on learning in the classroom, how teacher efficacy and growth is affected by Intech, and perceived changes in teacher development through InTech. Valuable comments about improving InTech for future training provide another theme. These themes yield an interesting corollary to the quantitative findings.

Quantitative Methodology

Twenty-five student teachers replied to both the pre- and post-training questionnaire thus providing an adequate initial sample (Sample group came from Fall of 1999). The questionnaire asks respondents to rate themselves on their ability to use equipment effectively in classroom instruction. Both the pre-test and post-test use a five-point scale for scoring answers (1=not comfortable at all, 2=slightly more comfortable, etc., and 5= comfortable enough to use in classroom). Student's pre-test answers are then compared with their post-test answers. Next, the differences in responses for each question between the pre- and post-test are compared for statistical significance using a paired sample t-test. A standard confidence interval of 95% will be used to determine whether the differences are statistically significant.

Qualitative Results

Evaluation data indicate a positive response to the InTech training. Questionnaires, journal entries, e-mail responses, and open-ended survey responses indicate that teachers have grown in their use of technology in their classrooms. Teachers believe that they are better able to use technology to enhance both teaching and learning. Comments regarding the benefits of InTech include: "not afraid to experiment, overcame my apprehension, feel empowered to use technology, and overcame feelings of inadequacy, helped my self-esteem". Others responded: "learned how to integrate, life changing experience, format very effective, enabled new interactions and connections with my students, and should be required for all GA teachers, parapros, and administrators". Teachers felt that the required lessons and the sharing process encouraged them to implement what they had learned in training. One teacher commented: "This course has been an absolutely wonderful opportunity to grow both personally and professionally. It has challenged me to learn new technologies and to learn more about myself as a teacher and as a learner. I have grown to understand my strengths and weaknesses and how the utilization of technology can help my students accomplish higher levels of learning and 'better' learning."

Relevance

Most preservice teachers desire experiences in their teacher preparation that they feel are timely and relevant to their professional development. Two areas of particular importance to the preservice teachers in this program include the marketability of their InTech training and the importance of technology in our society today.

Foremost on the minds of many preservice teachers completing their teacher preparation programs is the process of getting a job. A number of preservice teachers expressed the idea that their resumes improved dramatically because of being InTech certified or the importance of teachers being technology savvy as we enter the 21st century.

I did learn a lot and felt this class was a valuable experience. I know that I, as a teacher, will be better prepared because of it. Not to mention the fact that everyone that reads my resume mentions with enthusiasm "Oh! You are in-tech certified!"

This class is important to me due to the fact that school districts look favorably on the completion of this course. This will be a very important fact as I look for a job.

I know that being InTech certified is very appealing to those who are looking to hire. I never realized how good that looked on a resume until I had several people complimenting me on taking the initiative. So, job wise, InTech is a definite plus to me.

My father is an elementary principal, and he told me that Intech certification is something he highly recommends and looks for when he is hiring. So, I definitely feel this class was worthwhile.

As you all mentioned, it is going to be very important to be familiar with technology as our society becomes more technology oriented.

This program is good and I feel it is a vital part of teacher training due to the direction technology is taking in our student's future.

Finally, of local importance to the preservice teachers in Georgia was a mandate by the governor of the state requiring all teachers to have proficiency in technology, either through InTech training or by passing a computer skills test. One participant stated: "I enjoyed the class and I feel like it will really help all of us in the future, especially considering Governor Barnes and his new education initiative."

Hands-on Learning

A number of preservice teachers valued the hands-on learning inherent in the InTech training. They felt that practicing various activities in the lab setting helped them transfer their knowledge to the classroom in a more practical way. One student remarked: "I believe this class is great and really teaches one a lot to take into the classroom with them." Another preservice teacher explained: "I have learned tons of ideas that will be helpful when I teach. It is very helpful to anyone in the teaching profession. You can take everything we have learned in here and suit it to what you are teaching."

One preservice teacher went so far as to suggest that learning a specific skill such as creating a Web folder on the desktop would benefit her more than any other information she had received in the teacher preparation program: "I think that learning how to create a web folder on the desktop was one of the most useful things I have learned in all of my education classes at UGA."

Finally, this preservice teacher expressed the importance of practicing with particular software programs and then utilizing them in her technology-integrated lesson plans: "I know more about specific programs and lessons I can use -- that I have already had the opportunity to use while student teaching -- and I have had some practice in implementing these lessons."

Clearly, the methods used to teach these preservice teachers allowed them to gain practical experience and apply that experience to the classroom.

Efficacy/Growth

The most prevalent theme contained in the preservice teacher free responses about their InTech training address their increased efficacy toward using technology in the classroom and their appreciation for the professional growth they experienced during the training. "Before this class, I had no idea how to integrate technology in the classroom, but now I have a variety of creative lessons in which technology is the key."

Over half of the preservice teachers expressed their increased efficacy toward integrating technology into the curriculum as a result of InTech training. They felt both more comfortable and more confident in their newly refined technology skills:

> I think the most valuable aspect of the training in this class was the confidence that I built to use technology in the classroom. Before I was scared to use computers and didn't really know what I could do with them. Now that I have been exposed to different software, programs, and ideas, I feel more confident to use them with my future students.

> It is the first technology integration course I've taken that has truly tried to change the attitude of using technology for "presentations" instead of integrating it into your lessons. It has truly made me think more about how I want to use technology in my classroom as a motivator and a resource to simply make the learning process more meaningful and enjoyable.

> I really have had a wonderful experience with the InTech program. I feel that I have much more confidence with computers and feel I am capable of implementing technology lessons in the classroom, with a feeling of comfort.

> I feel much more comfortable and confident with using technology in my classroom now compared to the beginning of the course. I have seen just how much the children love to work with it as well.

> I had a hard time deciding if I wanted to take this class, but I am so thankful I did. I think the main reason I did not want to is because I was not very comfortable with integrating technology into my lessons. I feel I have learned so much and have really gained confidence in using technology.

Others emphasized the knowledge that they gained while several documented the change in their teaching development due to their experience with In Tech.

I feel that when I become a teacher I will have a large bank of knowledge about technology integration to pool from and I will be able to offer my students fun and meaningful lessons using technology.

The second is the actual attitude change I gained about using technology in the classroom. In any past courses with technology I've had, we have learned how to make technology work for us, the teacher, so we could make our lessons or any presentations better.

However, no one has ever taken the approach of using technology *with* content to motivate students even further. This is not only a valuable lesson gained, but it is a big way that my teaching has changed.

I have learned to look for ways to integrate technology in my lessons where I normally would not have.

I felt it was very practical and helpful in clearing up areas I was not sure about and showed me useful knowledge of things to do with using technology that I did not know.

Reflecting back on my InTech experience I think I have a new perspective on technology in the classroom. I guess I just thought of computers being separate from the rest of the curriculum, instead of being able to be integrated throughout the subject areas.

InTech opened up a lot of different avenues that I could take with my students. I felt computer literate prior to coming to this class, but now I feel I can actually make the technology come to life for my students.

While only two preservice teachers made reference to InTech training and experienced teachers, their comments are useful to consider as we move forward in preparing tomorrow's teachers to integrate technology. One preservice teacher, who made up missed time during classes with inservice teachers felt fortunate to have their ideas to consider alongside her own: "Professionally I have learned a lot from simply trying to teach these lessons in a hectic schedule, and by listening to the experienced teachers I had the opportunity to work with on make-up days. I learned so much more by listening to their ideas and concerns."

Two other preservice teachers wished that their cooperating teacher had already gone through the training or that they had been able to attend together:

I do wish that my mentor teacher had been InTech trained or was going through training at the same time. I would then have the computer and the programs in the classroom, and the teacher might better understand what my goals were.

It would be helpful if preservice teachers were all required to take this. It would also have been much easier for me if I had been placed in a classroom where technology was used by the students often. I just felt that I did not receive a great deal of support from my mentor teacher which made it difficult.

I feel that in-tech was a valuable experience as a pre-service teacher. However, I had limitations based on the fact that I was in another's classroom.

These preservice teachers articulated the importance of the timing of the InTech training, preferring to receive the training before beginning her teaching career.

I am very glad that I had the opportunity to attend InTech training, especially at this time in my professional career which is prior to actually having my own classroom.

This was a great experience. It would be helpful if preservice teachers were all required to take this.

Clearly, the importance of experiencing InTech training during the student teaching period enhanced the preservice teachers' sense of professional growth. The InTech facilitators aided in this professional experience: "The instructors treated us as professionals which really made us feel like actual teachers."

Critiques

Though the majority of the preservice teachers expressed gratitude and appreciation for their InTech experience, a handful had suggestions for improving the experience for other preservice teachers. "The days were sometimes very long and a little boring. I wish we could have practiced the technology and programs in relation to lessons and units we were actually teaching."

Two of the preservice teachers expressed concerns over the ability to transfer the experience of particular software applications to other more available technology. Interestingly, the notebook accompanying the InTech training contains specific information on how to perform the various InTech activities using alternative software programs.

I really liked InTech but I have a couple of concerns. Now that we know how to do all these programs, we will not have access to them.

The school I was placed at had limited resources, so I had to improvise a great deal. It would be beneficial if you would provide hints for using programs like word, works, or excel to do some of these things since these programs are found on many computers.

One of the preservice teachers offered this suggestion to improve the InTech experience for others like her: "I really like having to use the rainforest for a topic of everything we did in class. It would have been more useful if we could have done the activities on the topic of our unit." During their student teaching, the preservice teachers are required to plan and implement a two-week thematic unit. Throughout the InTech training, the rainforest theme is used but several preservice teachers were indeed able to transfer these activities into technology-integrated lessons that enhanced their own thematic units.

Though numerous preservice teachers felt that InTech training belonged in the final field experience of student teaching, one preservice teacher disagreed: "I wish that it had not occurred during our student teaching, as it seemed to add one more responsibility to my day."

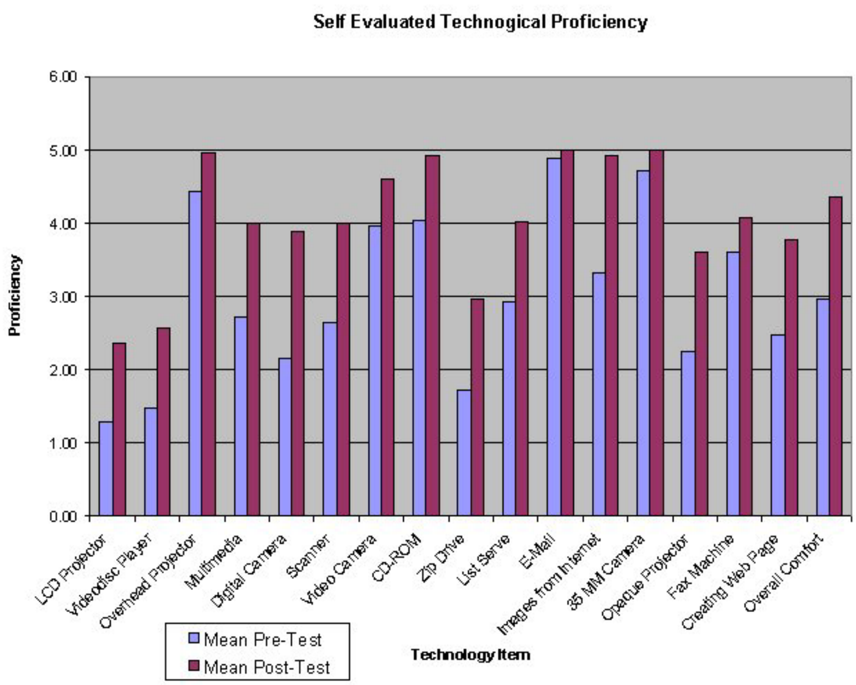
Quantitative Findings

Quantitative findings also give strong evidence that preservice teachers feel more technically competent after training. Respondents felt more competent on each piece of equipment after the InTech training. These positive gains are statistically significant across the board except for the gains in usage of e-mail and using fax machines. In these cases, prior experience limited the gains below a standard significance level. For example, the average pre-test score for e-mail usage was 4.88 and the post-test was 5.00. However, the gain was still positive. The largest gains reported by preservice teachers come from experiences less available in classrooms or home such as digital cameras, capturing images from the Internet, and on scanners and opaque projectors. Other substantial gains came in the creating of web-pages, using computer multimedia and zip drives, and using lcd projectors or videodisc players. Preservice teachers also registered an increase in being able to utilize list-serves in the classroom. Other, somewhat smaller, gains in student teacher proficiency include using overheads, video cameras, cd-roms, and 35mm cameras. However, these gains are positive and significant at the .05 level of confidence indicating that InTech training increased proficiency in these items as well. The following table is a breakdown of the statistical analysis for each item on the questionnaire.

Table 1. *Paired T-Test Differences*

Item Description	Mean Difference Pre and Post Test	S.E.	Lower Value	Upper Value	T-Score	Sig.
LCD Projector	1.08	0.28	0.50	1.66	3.83	0.001
Videodisc Player	1.08	0.24	0.58	1.58	4.42	0.000
Overhead Projector	0.52	0.18	0.14	0.90	2.83	0.009
Multimedia	1.28	0.27	0.72	1.83	4.78	0.000
Digital Camera	1.72	0.27	1.15	2.29	6.28	0.000
Scanner	1.36	0.34	0.67	2.05	4.05	0.000
Video Camera	0.64	0.26	0.11	1.17	2.49	0.020
CD-ROM	0.88	0.23	0.40	1.36	3.77	0.001
Zip Drive	1.24	0.35	0.51	1.97	3.52	0.002
List Serve	1.10	0.37	0.34	1.86	2.98	0.007
E-Mail	0.12	0.09	-0.06	0.30	1.37	0.185
Capturing images from the Internet	1.60	0.27	1.04	2.16	5.91	0.000
35 mm Camera	0.28	0.14	0.00	0.56	2.06	0.050
Opaque Projector	1.36	0.42	0.49	2.23	3.24	0.003
Fax machine	0.48	0.37	-0.29	1.25	1.28	0.212
Creating web page	1.30	0.35	0.58	2.02	3.70	0.001
Overall comfort level with technology	1.40	0.16	1.06	1.74	8.57	0.000
N=25					d.f.=24	
Overall Avg. Difference in Scores	1.03					

Figure 1. Comparison of Pre-Test and Post-Test Means



As illustrated by the table and graph above, the size of the positive increase in proficiency varied from 0.12 in e-mail usage to 1.72 for using a digital camera. Overall, the mean gain in proficiency was 1.03. The respondent's pre- and post-test comfort level response increased significantly from 2.96 to 4.36. These findings in general indicate that the InTech experience provided these preservice teachers with increased skills for integrating technology with classroom instruction. In particular, student teachers receive the greatest benefit of InTech training on items that they have little familiarity with using.

Summary

The quantitative and qualitative results indicate that student teachers feel much more competent to use technology in the classroom after attending InTech training. The student teachers indicated in surveys that InTech increased their ability to utilize specific items of technology in the classroom by statistically significant amounts. Student teachers register the largest increases in comfort levels for using items in classroom instruction on items that they have not experienced previously at home or class. Objectively, this sample of student teachers feels more competent to use items of technology in classroom instruction.

Subjectively, students for the most part expressed appreciation that they received this training before beginning their jobs as teachers. Others commented that they felt much more comfortable in being able to use technology as a valuable resource to integrate with their classroom instruction. These student teachers also felt that InTech training was a valuable job credential and that the hands-on training provided a greater sense of mastery of integrating technology into lessons. Thus, this sample of student teachers feels that InTech provides relevant, hands-on training that leads to a greater sense of efficacy in integrating technology into classroom instruction.

Further Work

Data collection and analysis should continue for additional cohorts of student teachers as well as their supervising teachers and professors. So far, the data gathered clearly indicate that teachers believe the InTech training and model to be effective. However, we do not have data that examines the degree of technology integration after training is completed. Additional data should shed light on the integration of technology into the curriculum not only among pre-service teachers, but among their mentors as well.

The InTech training now includes the issuing of a notebook computer with extensive software for each participant during his or her student teaching semester. Research needs to be conducted on the effect that the availability of this hardware and software has on the integration of technology.

This project was started with early childhood Preservice students, their cooperation teachers in the school systems and their supervising college professors. The program has grown and now includes agricultural education and social science education students. Both of these new groups are secondary education students and approach their teaching differently from the early childhood student teachers.

References

Eisenhower Professional Staff Development (Title II). *1995 Needs Assessment for the State of Georgia.*

Cooper J. M., & NCATE Task Force on Technology and Teacher Education (1997). Technology and the New Professional Teacher: Preparing for the 21st Century Classroom. *NCATE*, http://www.ncate.org/accred/projects/tech/tech-21.htm

Brooks, D., & Kopp, T. W (1990). Technology and teacher education. In W.R. Houston (Ed*.), Handbook of research on teacher education: A project of the Association of Teacher Educators* (pp. 498-513). New York: Macmillan.

Schrum, L., & Dehoney, J. (1998). Meeting the future: A teacher education program joins the information age*Journal of Technology and Teacher Education, 6(1),* 23-38.

Siegel, J. (1995, May/June). The state of teacher training. *Electronic Learning.*

Green, K. C., & Gilbert, S. W. (1995). Content, communications, productivity and the role of information technology in higher education. *Change*, 27(2), 8-18.

Smith, L. L. (1999). The Future of Technology in Teaching. *USA Today Magazine,* (March, pp. 26).

Relationships between Promoting Self-Regulatory Skills And Facilitating Student Interactions in Online Learning Environments

Yong-Chil Yang
Andong National University in Korea

Abstract

The use of self-regulatory skills (SRS) and student interactions have become as major research issues for successful and effective online learning. The purpose of this study was to investigate the effects of student interactions on promoting student SRS in online learning environments. Research questions are whether design strategies to facilitate student interactions have impacts on promoting the use of student SRS, or vice versa. It was expected that student interactions in the environments could contribute to the acquisition or use of their SRS. 32 college students were taught in online learning environment for 16 hours over 8 weeks. With 2 hours per week, one hour was assigned to students' content learning, and the other one for group discussions, having four members each group. SRS questionnaires were administered before and after online learning. Results from this study showed that embedded SRS strategies and group discussions in online learning environments have a possibility of contributing to the acquisition and use of their SRS. In the light of such findings, instructional design strategies to promote student SRS and to facilitate student interactions in online learning environments were discussed.

Introduction

It is well known that learners are responsible for their own learning. It means that "learning is not something that happens to students; it is something that happens by students"(Zimmerman, 1989, p. 21). To have this happen, it seems to be crucial that students have some extent of their self-regulatory skills (SRS) before learning. However, it is likely that most of public school students or many of college students would not come into learning settings with appropriate skills of regulating their own learning.

Recently, universities are moving more and more towards flexible modes of delivering courses. While high school has traditionally been a face-to-face experience, post secondary education is limited to its contact time and is being increasingly channeled through information and communications technology or Internet resources (Roblyer, & Wiencke, 2003). This fading of close social interactions in college significantly diminishes the regularly mechanisms that ensure students' natural development (McMahon, & Oliver, 2001; Chou, 2003). The lacks of close contact between instructors and learners or among learners prevent learners from developing their self-regulated learning. Lidner and Harris (1993) provided some evidence that self-regulated learning are an important part of academic success for college students. Peverly et al. (2003) showed that college students were not as good as expected at self-regulation. It means that many of college students need self-regulation training. But they had few opportunities to become self-regulated in their secondary school years, and as a result, they have few even if any SRS strategies (Orange, 1999). In particular, it has known that student SRS are required to be successful in online learning (Cennamo, & Ross, 2000; Cho, 2003). Accordingly, it seems that design strategies to promote the acquisition or use of student SRS are main research issues in online learning environments. Also, it is expected that design strategies to facilitate student interactions contribute to promoting acquisition or use of student (MacGregor, & Atkinson, 2002-03; Oliver, Omari, & Herrington, 1998).

Interaction seems to be an important factor in online learning environments as well as in face-to-face learning settings (Dykes, 2001; Hillman et al., 1994; MacGregor, & Atkinson, 2002-03). Especially, it is shown that learner interactions can become a good strategy for building learning environments that entails learners expressing their knowledge, getting feedback and constructing knowledge. More generally, learner interactions in online may be useful in light of pedagogical goals (Craig et al., 2000; Vrasidas, & McIsaac, 1999). Cazden(cited in Dytes, 2001) summarized the cognitive benefits of learner interactions: (1) learners are forced to confront each other; (2) learners can enact complementary roles, provide mutual guidance and support, and can service as scaffolding to help each other accomplish learning tasks that might otherwise be too difficult; (3) learners can find a direct relationship with a real audience from which they can get meaningful feedback; and (4) learners experiment and construct new understanding and ideas in peer discourse setting.

It is important that learners know when they need to ask help and be willing to ask for or to accept assistance from someone in learning settings. Orange(1999) suggests that such self-regulation as asking help, giving feedback, self-monitoring, making elaboration and organization can be taught and that peer models or cooperation can be effective means for teaching self-regulation. Therefore it is expected that student interactions in online learning environments contribute to the acquisition or use of their SRS. More recently, researchers reported that promoting students' SRS is possible by their own efforts such as student interactions or cooperation (Craig et al., 2000; Hartley, 2001; Orange, 1999). If then, it appears significant to investigate the effects of student interactions on promoting the acquisition or use of his/her self-regulatory skills in online learning environments.

The purpose of this study was to investigate the relationship between promoting student interactions and facilitating SRS in online learning environments. In order to explore any relationship between these two, design strategies to embed some components of SRS and student interactions through discussions need to be applied to authentic online learning environments. It is assumed that the degree of promoting the acquisition or use of learner SRS might determine the amount and depth of his/her learning in online learning environments, especially when designing web-based learning environments.

Methods

Participants

32 undergraduate students from one university in Korea participated in the study. Participants were enrolled in a two-credit required course that is Introduction to Educational Technology. There were 8 males and 24 females, all junior students with ages ranging from 21 to 24, respectively. They were divided into 8 groups, and each group has 4 members. They were randomly assigned to each group because they are comparatively familiar with each other to induce their active participation at group discussions

Materials

The course for this study was developed into online learning materials. It was taught under online learning environments. The amount of learning materials is for eight weeks with 2 contact hours per week. Participants logged in on the site and learned the course once a week for 8 weeks. Some of strategy components to promote SRS during learning from online were embedded into the contents of the course.

The learning activities in each lesson were designed to encourage learners to study the materials in online and to enable them to deeply explore a given topic through group discussions every class meeting. The order of the presentation was lesson objectives, outlines, content, summary, preview of the next lesson, and discussion about given topics. There were also five quizzes included in each lesson. Discussion between group members was made for 50 minutes, and then each group submitted its result to the discussion task room in an online environment.

Embedded SRS Strategy Components

Strategies to Promote The Use of SRS

Strategies for promoting learner SRS were embedded into the learning materials. SRS strategy components employed in this study were performance control strategies (self-instruction & self-monitoring), self-efficacy strategies (peer feedbacks & attribution feedbacks), and cognitive strategies (elaboration & organization). These SRS strategies were selected because learners have hardly acquired or comparatively used them less than any other strategies during their learning. The embedded SRS strategies in learning materials are the following:

Performance Control Strategies (Zimmerman, 1998):

(a) Self-instruction: Telling for themselves in mind to memorize major concepts or key principles. Or students learned by reciting important concepts for themselves.

(b) Self-monitoring: Checking out for themselves the results of their own learning activities

Self-efficacy Strategies (Zhang et al., 2001)

(a) Peer feedbacks: Encouraging or praise each other about peer ideas or comments from group discussion, or active participation during discussion.

(b) Attribution feedbacks: Providing effort feedbacks for prior success supports students' perceptions of their progress, sustains motivation, and increases self-efficacy for learning (Schunk, 2001). They were encouraged to give attribution feedbacks to each other during discussion.

Cognitive Strategies (Park, 2000; Cho, 2003):

(a) Elaboration: Paraphrasing with learners' own words, and link new information to what they already know when learning key words or concepts in the text.

(b) Organization: Summarizing the contrast, and developing a table or a figure mentally. They were asked to summarize their learning materials on the provided board.

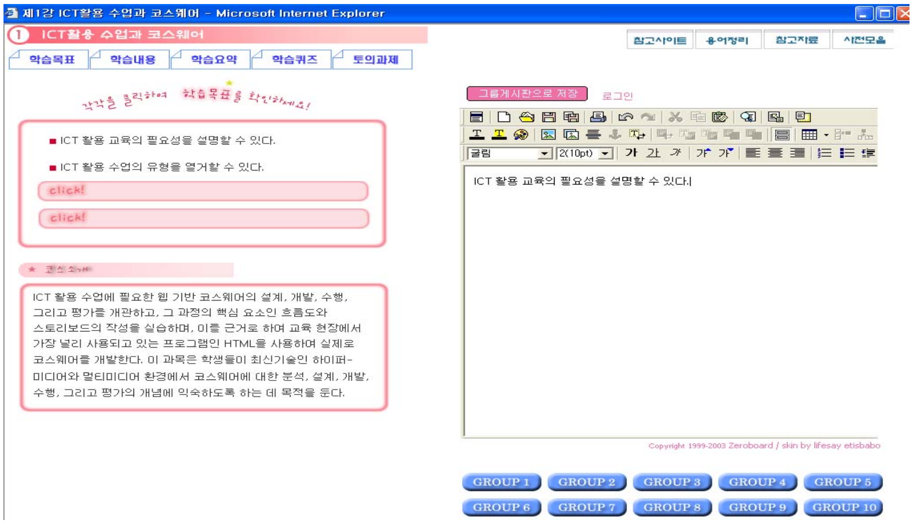


Figure 1 Text screen with embedded SRS strategies

B. Strategies to Facilitate Student Interactions

(a) It focused on group communications through synchronous discussion. Each group selected one of the given topics through discussion.

(b) Student interactions were performed through discussion to explain and solve the topic selected by the group members for around 50 minutes per week.

(c) Discussion was synchronously made through the MSN messengers in an individual computer screen. Group members logged in after learning the tasks in online.

(c) Group members were encouraged to use attribution feedbacks to each other as often as possible during discussion, for instance, using expressions like "You work very hard!" "You're very knowledgeable on that issue!" when they have good idea during discussion. Also, the use of praise or encouragement words like "good job" or "very well" was encouraged among group members. And feedback examples were exhibited during discussion sessions.

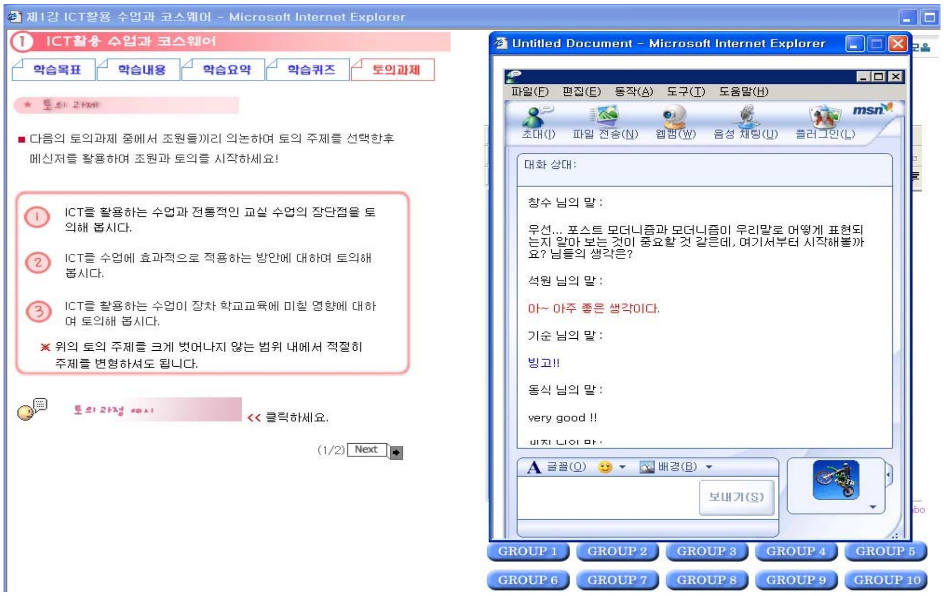


Figure 2. An online discussion screen

Measuring Instruments

SRS questionnaires adapted by Yang (2000) and Park (2000) from the original versions of Zimmerman and Martinez-Pons (1986), and Pintrich and De Groot (1990) were used to measure the acquisition of student SRS before and after online learning. The 84 item self-report measure addressed participants' SRS. Among the 84 items, items of performance control strategies were 14, self-efficacy strategies 11, and cognitive strategies 13. The attitude survey was composed of 15 items, and was administered right after learning. The frequencies and features of interactions between learners and the results of discussion on the given topics were collected.

Research Design

Research design used in this study was pretest-posttest control group design. While this design is one of the experimental designs used extensively when applying new methods or programs in education to instruction, it falls far short of handling the other sources of external validity (Tuckman, 1999). In the present study, the design was adopted because of exploring the effects of online learning programs on the acquisition or use of SRS through student interactions. The independent variables in the study were the embedded SRS strategies and student interactions during online learning and the dependent variables were scores yielded from student SRS measure, and responses to the attitude survey.

Procedures

The study was conducted as a part of students' regular classes for pre-service teachers training. The participants completed the SRS measure as a pretest right before the beginning of the study. They learned the course contents in an online environment for 8 weeks with two contact hours per week. Learning activities were divided into two parts. The first part was the online learning activities and the second was discussion activities among group members about given topics through the MSN messenger. Immediately after an eight-week study schedule, they took the SRS measure as a posttest, and the attitude survey.

Data Analysis

Data yielded from SRS measure were analyzed into mean scores on the pretest and posttest between the experimental and control group. The t-test was used, and the level of α set at .05. Results of the attitude survey were analyzed by contrast with SRS results. And interaction processes by each discussion group were examined.

Results and Discussions

The Effects of Embedded SRS Strategy Components

The effects of embedded SRS strategy components and student interactions through discussion on the acquisition or use of SRS were examined in online learning environments. To determine whether there was a difference between the pretest and posttest scores yielded from SRS measure on the experimental group, a t-test was used. As shown in Table 1, the mean difference between pretest scores (M = 3.44, SD = .46) and posttest scores (M = 3.60, SD = .44) in the experimental group was not statistically significant, t (31) = 1.669, $\underline{p} > .05$. This means that the use of embedded SRS strategy components and student interactions in online learning environment could not be useful in facilitating the acquisition or use of student SRS in this study.

Table 1 *Means and SD for Scores of embedded-SRS strategy components*

Test	n	M	SD	t	df	p
Pretest	32	3.44	46	1.669	31	.105
Posttest	32	3.60	44			

The Effects of The Whole SRL Measure

Although not statistically significant, the embedded SRS strategies and student interactions in online learning did appear to have substantially positive impact on the acquisition or use of student SRS as compared to the result of the control group. As shown in Table 2, the mean difference between posttest scores of the experimental and control group (M = 3.60 and 3.57) was far smaller than the mean between pretest and posttest means (M = 3.44 and 3.60) on the experimental group. This means that the use of embedded SRS strategies and student interactions in online learning could contribute to promoting the acquisition or use of student SRS (Cennamo & Ross, 2000; Hartley, 2001; Peverly et al., 2003), or to facilitating student interactions through the use of his/her SRS (Dykes, 2001; Orange, 1999).

Table 2 *Means and SD for the total SRS strategy scores by each group*

Group	n	M	SD	t	df	p
Experimental	32	3.57	.44	.419	65	.677
Control	35	3.53	.37			

The Results by Each of the Embedded SRS Strategy Components

It appears that the results were mainly yielded from group discussions in online learning because the mean difference between pretest scores and posttest scores for performance control strategies was larger than other embedded SRS strategy components (see Table 3). This means that interactions among students might be an important factor influencing the acquisition or use of SRS strategies. When they discussed about a topic with group members through online MSN messengers, 78% of participants (32) said that student interactions were useful to promoting self-monitoring during learning. Although both the use of self-efficacy strategies and cognitive strategies were improved during online learning and student interactions, none of them was statistically significant in the degree of improvement.

Table 3 *Means and SD by each of embedded SRS strategy components in online learning*

		Pretest		Posttest	
	n	M	SD	M	SD
Performance Control Strategies (Self-instruction & self-monitoring)	32	3.36	.67	3.59	.59
Self-Efficacy Strategies (Peer feedbacks: immediate and attribution)	32	3.34	.49	3.47	.48
Cognitive Strategies (Elaboration & organization)	32	3.47	.55	3.67	.54

Relationships of SRS Strategies and Student Interactions in Online Learning

On the other hand, the mean scores of the whole SRS strategies on the pretest and posttest of the experimental group are likely to show that the possibility of facilitating student SRS strategies through group interactions in online learning is much higher. That is, the mean scores of the whole SRS strategies on the pretest and posttest were further lower than mean differences on the acquisition or use of embedded SRS strategies on the experimental and control group (see Table 1 and Table 4). This suggests that student interactions through discussions among group members through the MSN messengers (student interactions) could influence promoting the use of student SRS strategies (McMahon, & & Oliver, 2001; Zhang et al., 2001), or that the acquisition or use of student SRS strategies might have an impact on facilitating student interactions in online learning.

Table 4 *Means and SD for the whole SRS on the experimental group*

Test	n	M	SD	t	df	p
Pretest	32	3.51	.40	.630	31	.533
Posttest	32	3.57	.41			

However, it is recognized that this study had several limitations, which could not gain more positive impact on the acquisition or use of student SRS strategies in this study. For instance, the study had comparatively short periods in online learning (McMahon & Oliver, 2001), an inappropriateness of embedded SRS strategies design (Cennamo & Ross, 2000), and the weak validity of SRS questionnaire to be measured in online learning (Orange, 1999). However, the results of the present study suggest that the acquisition or use of student SRS strategies in online learning environments can be facilitated through embedded strategies in lesson materials and student interactions using synchronous discussions. Further research on promoting the use of student SRS or facilitating student interactions is required to examine the effectiveness of each components of self-regulatory skills in online learning environments.

Conclusions

This study was conducted to investigate the effects of embedded self-regulatory skills strategies and student interactions through group discussions in online learning environments. The results from this study showed that the embedded SRS of performance control strategies, self-efficacy strategies, and cognitive strategies did not influence the acquisition or use of student SRS in online learning. Those strategies suggest that when compared to the whole SRS strategies, however, they could promote using student SRS. This means that promoting the use of SRS might be facilitated by student interactions through online group discussions. This study seems to have some implications for providing a possibility of exploring relationships between promoting SRS and facilitating student interactions which are necessary in designing and developing an effective online learning environment.

References

Berge, Z. L. (1998). Guiding principles in web-based instructional design.. Educational Media International, 35(2), 72~76.

Berge, Z. L. (1999). Interaction in post-secondary web-based learning. Educational Technology, 39(1), 5~11.

Cennamo, K., & Ross, J. (2000). Strategies to support self-directed learning in a web-based course. Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA. April 24~28, 2000.

Chou, C. (2003). Interactivity and interactive functions in web-based learning systems: A technical framework for designers. *British Journal of Educational Technology*, 34(3), 265~279.

Cho, M-H. (2003). The effects of design strategies for promoting self-regulated learning on students' self-regulated learning skills and achievements in WBL environments. Unpublished thesis. Andong National University, Andong, Korea

Craig, D., ul-Hag, S., Khan, s., Zimring, C., Rick, J., & Guzdial, M. (2000). Using an unstructured collaboration tool to support peer interaction in large college classes. In B. Fishman & S. O'Connor-Divelbiss(Eds.), *Fourth International Conference of the Learning Science*(pp. 178~184). Mahwah, NJ: Erlbaum.

Dykes, M. (2001). Assessment and evaluation peer interaction using computer-mediated communication in post-secondary academic education. http://www.usask.ca/education/coursework/802papers/dykes/dykes.htm

Hartley, K. (2001). Learning strategies and hypermedia instruction. *Journal of Educational Multimedia and Hypermedia*, 10(3), 285~305.

Hillman, D., Willis, D., & Gunawardena, C. (1994). Learner-interface interaction in distance education: An extension of contemporary models and strategies for practitioners. *The American Journal of Distance Education*, 8(2), 29~42.

Lidner, R., & Harris, B. (1993). Teaching self-regulated learning strategies. Paper presented at the convention of the AECT, New Orleans, LA.

MacGregor, S., & Atkinson, T. (2002-03). Facilitating learner interaction in the two-way video classroom. *Journal of educational Technology Systems*, 31(1), 45~62.

McMahon, M., & & Oliver, R. (2001). Promoting self-regulated learning in an on-line environment. In ED-Media 2001 World Conference on Educational Multimedia, Hypermedia & Telecommunications. Proceedings (13th), Tampere, Finland, June 25~30, 2001

Oliver, R., Omari, A., & Herrington, J. (1998). Exploring student interactions in collaborative world wide web computer-based learning environments. *Journal of Educational Multimedia and Hypermedia*, 7(2), 263~287.

Orange, C. (1999). Using peer modeling to teach self-regulation. *The Journal of Experimental Education*, 68(1), 21~39.

Park, H-K. (2000). The effects of different ways of employing self-regulated learning strategies in CBI. Unpublished dissertation. Florida State University, Tallahassee, FL

Peverly, S., Brobst, K., Graham, M., & Shaw, R. (2003). College adults are not good at self-regulation: A study on the relationship of self-regulation, note-taking, and test taking. *Journal of Educational Psychology*, 95(2), 335~346

Pintrich, P., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 23~40.

Roblyer, M., & Wiencke, W. (2003). Design and use of a rubric to assess and encourage interactive qualities in distance courses. *The American Journal of Distance Education*, 17(2), 77~98.

Shin, 1998

Schunk, D. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman, & D. H. Schunk(Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives*(2nd ed., pp. 191~226). Mahwah, NJ: LEA

Tuckman, B. (1999). *Conducting educational research*(5th ed.). Forth Worth, TX: Harcourt Brace

Vrasidas, C., & McIsaac, M. (1999). Factors influencing interaction in an online course. The *American Journal of Distance Education*, 13(3), 22~36.

Yang, M. H. (2000). The study on development and validation of a self-regulated learning model. Unpublished doctoral dissertation, Seoul National University, Seoul

Zhang, J., Li, F., Duan, C., & Wu, G. (2001). Research on self-efficacy of distance learning and its influence to learners' attainments. Unpublished research paper. Tsinghua University, Beijing, China.

Zimmerman, B. J. (1989). Social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, 81, 329~339.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk, & B. J. Zimmerman(Eds.), *Self-regulated learning from teaching to self-reflective practice*(pp. 1~19). New York: Guilford Press

Zimmerman, B., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614~628.

From Outcomes to Process: Exploring Students' Lived Experience of Web-based Learning

Li Zhu
Ian Baptiste
The Pennsylvania State University

Abstract

The purpose of this qualitative study was to describe students' lived experiences of Web-based Learning (WBL). Observation techniques and in-depth personal interviews were used to collect data. The findings of this research were grouped into three themes: opportunities and challenges, rewards and benefits, and evaluation of instructions and learning. The research contributes to the discourse on web-based learning in the following areas: accessibility, interactivity and instructional strategies.

Introduction

There are numerous names for web-based learning (WBL): Web-based Instruction (WBI), Internet-based Training (IBT), Web-based Training (WBT), Distributed learning (DL), etc. Jolliffe, Ritter & Stevens (2001) defined web-based learning as "…the delivery of and access to a co-ordinated collection of learning materials over an electronic medium using a Web server to deliver the materials, a Web browser to access them and the TDC/IP and HTTP protocols to mediate the exchange" (p.8).

Khan (1997) defined Web-based learning as "a hypermedia-based instructional program which utilizes the attributes and resources of the World Wide Web to create a meaningful learning environment where learning is fostered and supported" (p. 6). He asked the question "what does it take to provide the best and most meaningful flexible learning environment for learners worldwide?" and developed a framework for creating meaningful distributed learning environment which include eight dimensions: pedagogical, technological, interface design, evaluation, management, resource support, ethical, and institutional.

Jung (2001) maintained that although some studies tried to connect the technical features of the Web to the problems of designing web-based learning, such as Khan (1997), Hiltz (1994), Simich-Dudgeon (1998), and Barron (1998), all these studies showed "little linkage to established pedagogical theory in general or to distance education theory in particular" (p.525). After reviewing fifty-eight articles from six refereed international journals in the fields of distance education and educational technology, he found out that not many studies paid attention to the pedagogical processes in WBL, i.e. what was really happening in teaching and learning processes of WBL and why it happened. Instead, the most frequently asked research questions focused on how to design effective WBL, how to encourage interaction, and what were the effects of WBL on learner satisfaction and perceived learning outcomes. In particular, he pointed out that there is little research that examines the characteristics of distance education in the Web environment (Jung, 2001).

Research Purpose and Question

The purpose of this study is to describe students' lived experiences of web-based learning. It examines the very nature of the web-based learning phenomenon. Understanding the students' lived experience of web-based learning can facilitate the design and development of an effective and efficient web-based learning environment.

The major research question is: How do students enrolled in a WBL class describe their experiences and what meanings do they attach to those descriptions?

Methods

Participants

The participants of this phenomenological study were 30 students enrolled in the Nutrition 360 class in Fall, 2002 in large land grant university in northeastern USA. Most of the students were sophomore or junior students majoring in Nutrition and two of them were from other programs.

Eight students who took this class before in either Fall, 2002 or Spring, 2003 accepted the invitation to be interviewed for one-hour. Seven of them are either junior or senior students majoring in Nutrition who didn't have any previous experience in web-based learning. They are all female Caucasians. The only male student is an African-American/native American who has a B.S. in Kinesiology and a minor in Nutrition. He had extensive knowledge with computers and technology before he took this class and is currently working in the media technology field. He also had previous web-based learning experience before. Most of the interviewees consider themselves as hands-on type of learners who prefer to learn things by actually doing them. One student said she is a very visual person. Another one said she learns well when working in groups.

The Class

The Nutrition 360 class is a traditional face-to face class that integrates the web-based learning instructions as part of the class in order to achieve the course objectives. One of the objectives of the class is that the students will be able to create a personal web site as well as a professional online portfolio as a part of their web site. Therefore, WBL instructions were created to help students develop their web site and online portfolio. Although a technology TA spent several class sessions walking them through the steps, the students needed to spend a lot of time going through the WBL instruction by themselves in order to accomplish the projects.

Research Design and Rationale

This is a phenomenological research. Phenomenology is chosen as a major research method based on the research purposes and questions. First, phenomenological research is primarily concerned with understanding and describing people's lived experiences. Phenomenology asks the simple question, what it is like to have a certain experience. The purpose is to gain a deeper understanding and insightful descriptions of the nature or meaning of our every experience (Van Manen, 1997). Second, phenomenological research is the study of essences of a phenomenon. It "asks for the very nature of a phenomenon, for that which makes a some-"thing" what it is-and without which it could not be what it is" (Van Manen, 1997, p.10). It searches for the invariant structures of individual's experiences. Third, phenomenological research is the description of the experiential meanings we live as we live them. It "attempts to explicate the meanings as we live them in our everyday existence, our life world" (Van Manen, 1997, p.10). Based on the research method, the strategies/tactics the researcher utilized to collect data included close observations and in-depth interviews.

Procedures

Observation. The researcher conducted directly observation of the phenomena as a participant during six WBL class sessions in different computer labs throughout the semester. The main objective for the six WBL class sessions was to teach the students how to create a web site that includes an online portfolio. The students were first showed how to create a web site and an online portfolio. Then they practiced the skills on their own with the help of the web-based instructions on the course web site. The researcher conducted direct observation by walking around the classroom observing the students in the WBL setting. Detailed experiential anecdotes were kept to report what exactly happened that made a situation stand out for the purpose of the study.

In-depth interviews. A single one-hour in-depth individual interview was conducted with eight participants. The researcher selected the participants on a voluntary basis by describing the purpose of the research and encouraging them to accept the interview. An interview protocol with open-ended questions was designed as a guideline to conduct the interview. The students were encouraged to explore the whole WBL experience.

Findings

The researcher grouped the students' lived experiences under three major themes: opportunities and challenges, rewards and benefits, and evaluation of instructions and learning. Three general questions were raised according to the three major themes: 1. What opportunities / challenges attributed to WBL do they describe? 2. What rewards and benefits attributed to WBL do they describe? How do they evaluate the instructions as well as their own learning? Each of these themes, and their relationship to each other are now discussed.

Opportunities and Challenges Attributed to the WBL

Web-based learning, as a comparatively new learning environment, has created opportunities as well as challenges for the students. Participants noted that WBL, compared to traditional learning, is very accessible in various dimensions. This accessibility characteristic of WBL, in turn, motivates them to learn.

The participants (from different perspectives) described various dimensions of the accessibility of WBL. They suggested WBL is very accessible because of the following reasons: It's open to anyone with an Internet connection and the information is always there on the web site. Therefore, they can always go back and review things whenever they need or whenever they feel like it. They can also save their work where they are stuck and go back and try it later. Compared to getting books from the library, it's convenient to look up and do things in WBL. Because the course web site is well organized, it's easier to navigate and locate information they need at a certain time. They always have the option if they want to go beyond the class requirements and learn more. JL2 said,

> But the way we did it, I took my web page to the [web site] and I put every pictures of myself. I really developed into the portfolio. I could do it if I wanted to. And I can continue using it adding to until I graduate and after. So you can work within your means. That was really nice too. So every one will get a grade and fulfill certain requirement. But you can go beyond that if you wanna learn more. That was a good option.

CY1 provided another dimension of accessibility of WBL. He suggested that sitting in front of the Internet is like being in class because it's as if there is "somebody" there to present the information to you. He said,

> Web-based learning because it's so accessible, there is even...I think… wbt.cac.psu.edu….has a whole list of web-based learning stuff. And because it is accessible, which is another key. It made me want to learn things that maybe I didn't want to learn before, maybe I didn't have time before. You know, Visual basic, or something like that. Ordinarily, I'll say forget it. If you throw me a book right in front of me and said: "I want you to read this cover to cover?" You might say "Yeah! " but when you talking about sitting in front of the Internet, having somebody present it to you, almost as like being in class. It's a little more compelling. I feel a little bit more incline to learn from a web-based medium than text… I would much rather read or learn from web-based medium than traditional text. It's like I would rather read an e-book rather than a traditional text. I'll go to the library any day before downloading an e-book, or for whatever reason.

This "somebody" seems like a personal tutor who makes CY1 feel more incline to learn and want to learn some things that maybe he didn't want to learn or didn't have time to learn before. It suggests that one of the dimensions of accessibility of WBL—like having "somebody" present it to you—motivates students to learn.

In a word, WBL provides students with expanded learning opportunities through different dimensions of accessibility characteristic, which in turn, serves as a motivation of students learning in the WBL environment.

In addition to opportunities, most participants also mentioned all kinds of challenges they face in the WBL environment. These challenges fall into three categories: not familiar with the technology, lack of curricular and technical support and the emotional disturbance resulted from these challenges.

Unfamiliar with the technology. Seven participants were not very familiar with technology. Although they did have some knowledge with computers from previous education, it was their first experience in a WBL environment and it was also their first time to do a big online project—creating their personal web site as well as their online portfolios—using web-based instructions. They were not used to this new learning environment full of technology. JL2 told her experience as below,

> I'm not very computer savvy. And I find that with the most things I have to learn there is some answer to it, or you have an opinion about it, or something, like, I mean, there is a certain methodology to it. But I understand what the computers are too, but it was also new to me that like you tell me to do something but then I go and do it, it come out different, I freaked out and I didn't know the answer, how to fix it.

Lack of curricular and technical support. Because of the fact that for most of the students, it was their first time to go through WBL, they especially needed technical and curriculum support, and in most of the situations, real time support. All the students interviewed suggested they needed either curriculum or technical support to certain extent when they went through the web-based instruction.

Four students specifically pointed out that they couldn't' ask questions real time in WBL environment. CY1 branched out a little bit by including his WBL experiences other than the Nutrition class. He described a specific time, in the WBL environment, that he felt he couldn't' ask questions real time when he wanted to.

> I tried to learn visual basic through Web-based learning solely. I really suffered from the exact problem I just said, I couldn't ask questions. We are talking about writing code. The laws are something… is extremely detailed. You can do this, but not this, it's case-sensitive, it's not case-sensitive, you need a comma here… the syntax here because of this and you know I was trying this on my own, and sometimes I get errors. And I would very much like to say to a teacher look at this code, I think it's right, you know, whatever, how come it is incorrect, for whatever reason. And I really didn't have the opportunity to do it. Because not that interactive yet. I really want to learn about my own mistakes as I was trying it, and like I said, eventually it will be there, but it's just not there yet.

JR4 also described the same problem she had with WBL,

> It's just frustration. Umm… I basically just give up the web site and went to ask some one else. That's the only problem with the web site that you can't ask direct questions on the web site. You can always ask someone else in your class or ask the teachers.

Lack of empathic, real time feedback appears to be another problem with WBL. CY1 said he wanted to ask questions and get feedback while he was doing WBL. However, he suggested, nobody was there to look at and experience what he was doing. Therefore, nobody could understand exactly what he was going through.

> I wanted to ask questions, and I wanted feedback. I wanted to know what I was doing wrong. You know, I could not get feedback, I could not learn about my own mistakes because nobody is looking at that, nobody can see that, nobody can experience that as I am doing that.

CY1 further stated that in WBL, it was very hard for the teachers to understand his problems or questions on the fly. His question was very specific and came from a specific situation as well. The teacher might not understand exactly what he means, therefore, couldn't give correct feedback.

> Right now, there is sort of a problem with interaction in web-based learning. If I asked a…question. For example, something very detailed whatever. It's very hard for the teacher to understand it on the fly. I wanna ask something very specific, and then maybe they wanna to go another angle with it. maybe they thought I meant the wrong thing. It's very hard to deal with that and say well, that's not really what I meant, or whatever. I mean you can not ask questions. It's getting there.

Three participants suggested that they need interactions. They want people to interact with them, or to explain it fully what it means. They want personal communications. And when they do it in class, they want more people to float around helping them.

Emotional Challenge. Because of their unfamiliarity with the WBL environment and the lack of the curriculum and technical support, the students were enduring all kinds of emotional challenge when they were going through the WBL. The following feelings were described by the participants: frustrated, confused, stressed, insecure, vulnerable, angry, disappointed, intimidated, scared, and unfulfilled.

Most of the students felt frustrated because one or more of the following reasons: couldn't follow instruction; couldn't get help; didn't know what to do; it was hard to do project online; it was hard to find things online; had to go back and forth and have multiple windows open; might lost the work you've done; sometimes, you were just doing it, but didn't know why you are doing it; when other students knew what's going on, but you didn't; things didn't work and had to wait for help.

JL2 felt vulnerable and insecure because WBL was not something she was used to.

> I just think that there is a lot of insecurity. You felt more vulnerable than secure because there is something you couldn't figure it out unless spend time with it or get help or just try it again. I think that.. that's not what we are used to, we are used to book learning, there is a right answer, take this test… I'm in this huge pool of information, and I had really a lot to learn…there are so much I

didn't know…again, insecurity I guess.

Not only herself, JL2 also found her classmates felt vulnerable and insecure as well, and she was disappointed by her classmates' reactions to the vulnerable and insecure feelings they were going through.

> I really love doing that kind of stuff, but I really found that the most the people in my major don't like those stuff. It's just the people in that class because they feel so vulnerable and insecure that they just like projected everywhere else.

Rewards and Benefits

This theme describes the tangible and intangible rewards and benefits brought by WBL despite all the challenges they face in the WBL environment.

Tangible benefits. The tangible rewards and benefits described by the participants fall into three categories: help to achieve the course goal and objectives which include building a personal web site as well as a professional online portfolio; contribute to their future career and personal development, and facilitate the transfer of learning to other areas of study.

JL2 said her web site and the online portfolio helped her a lot when she applied for an internship. CY1 suggested WBL experience helped him transfer what he had learned to other platforms and areas.

> Same things, a lot about carrying over to other platforms and other areas. It allowed me to branch out a lot. Even with web creation… how to incorporate multimedia into web pages, for example, how to incorporate multimedia into applications, things like that. Things I didn't really think about before, even I didn't do that, I had the idea what the concept, I want to play around with it. I got curious from what I learned…go with that angle.

Intangible benefits. The intangible benefits described by the participants were emotional benefits they gained from the WBL experience. The emotional benefits include: their confidence is increased because they feel they can handle more; it's not as hard as they thought or it's easier to locate things; it's enjoyable to put stuff online; feel more open to other types of things; feel accomplished something which is pretty hard; feel enlightening; feel happiness; feel comfortable to know how things worked; feel compelled to learn; feel curious from what they have learned; feel elated when they put it right; feel excited when things work; feel proud when they figure it out on their own; feel relieved when they learn, etc.

Evaluation of Instructions and Learning

This theme includes the students' evaluation of the instructional strategies used in WBL as well as their evaluation of their own learning experience.

The participants mentioned various instructional strategies used in WBL. It includes the use of correspondent images, the use of instructional purposes and objectives, the use of step-by-step instruction, the organization and categorization of the content and learner control.

Use of correspondent images. Correspondent images here are defined as the screen shots serve as visual illustrations right next to the text descriptions. Three students mentioned that the use of correspondent images of the WBL was good. RA8 mentioned it made it easier for her to find instructions. HR7 said it reinforced what she needed to do in addition to the text.

> That helped… having the pictures because like some places you click on things, the works were pretty similar, you could have "Save as" this, you click on it…having the pictures right beside it made it easier to make sure you click on the right thing. I think having the visual help me out because I didn't know anything so there were two things reinforcing what I need to do.

CY1 said because every step of instruction was correspondent with images, it made it easier for trouble shooting. When he got stuck, he went back and looked at the instructions.

> The use of images is also good, you can tell where graphically as well as just the written part. If you go here, then you go here. The correspondent images, that was done very well. I admit I myself have to use it a few times when I got stuck. Great for trouble shooting…Your web site I used it mainly for trouble shooting. I'm stuck, then I went back on it and said ok, well you go here, here and there and then you become unstuck.

Use of instructional purposes and objectives. Several participants mentioned that they needed to understand the purposes and objectives before getting into the detailed steps. The purposes and objectives will give them an idea what they are supposed to do and why they are doing the steps. MT6 said,

> I guess make it more simple because there is so much information, it could get so complicated, and there is tons of information. But to fully explain it as a whole first, I know the web site does that…you need to explain it as a whole. For me, I need to understand the purpose the why, the objectives, and then say: "ok, well, since we are doing this, let's break it down to the first step."

Use of step-by-step instructions. Many participants indicated that step-by-step instruction helps. CY1 suggested because of the nature of the step-by-step instruction on the course web site, it was very easy for him to follow and especially good for people who don't have much technology background.

> Yours in particular is pretty good. It was very step-by-step, it was very easy to follow, and broke down to very fundamental level which is good because there are some people who look at anything technological, they just can't handle it for whatever reason. Yours is good in the respect that it had step-by-step, you press this button, then you press that button, you go to tools, then you go to internet options, this and that. It's very difficult to be confused on, it's very difficult to not understand what's going on in that regard.

JA4 also thought the step-by-step instruction was very simple and easy to do.

> I learned a lot about computers just from following step by step, like, I needed that, step-by-step, like saying go to this place and there. It was nice to have that right next to what I was doing.

Organization and Categorization of the Content. HR7 stated that the content on the web site need to be organized and categorized well so that people can easily find things they need.

> If it is organized well, you can have a lot of stuff at one time. But if it's like some of the links I would go to find things to add to my web site won't as organized, I got lost and I'll just click the whole thing away because it was too much stuff at once, and it wasn't organized. So I think web-based things need to be organized or categorized well so that you can find what you need. Otherwise, there is no point to go in there if you can't get anything.

Learner control. Both HR7 and RA8 suggested it was good to be able to take the WBL at their own pace versus having to follow the professor's pace in traditional classroom. As HR7 suggested, different people have different paces to learn things. WBL has this unique characteristic to allow learner access learning at their own paces.

> With web-based learning, you take it at your pace. Versus other learning it's kind of as the professor go, you go. Web-based learning, based on how the professor moves from thing to thing, there is a pace. But the amount of time, and the pace you take your time to spend on that was more up to yourself, which I think it's nice. Different people pick on certain things faster than other things. So it's more personalized in a way.

Another part of the theme involves students' evaluation of their own learning experience in the WBL environment. When talked about how they know they learned something in WBL, TA3 said,

> I felt like I learned something when I can just do it over again, or if I feel confident about it or if I accomplished the task I was trying to do. I can do it again or if I can relate it to my personal life…I'll try to learn it, and then if I can do it again later, then, yeah, I did learn something.

Discussions

Accessibility

Accessibility surfaced as one of the major opportunities the students have in WBL. Some literatures have talked about the accessibility issue in WBL environment, for example, Khan (2000) included "information accessibility" as a factor to help create a meaningful WBL environment, but he didn't define what is "information accessibility" nor did he talked about what does that mean to the learners. Most of the literatures refer to the accessibility of the WBL as "Web accessibility", which is defined by LeTourneau (2003) as "…anyone using any kind of Web browsing technology must be able to visit any site and get a full and complete understanding of the information as well as have the full and complete ability to interact with the site - if that is necessary." Very few of them have explored the real meaning of accessibility of WBL from the students' lived experience. By exploring the different dimensions of accessibility in WBL, this research contributes to the literature by finding out the various dimensions of meanings that the students attach to the accessibility characteristic of WBL. To the students, accessibility of WBL means it's open to anyone with an Internet connection, means being able to go back and review anytime and anywhere, means being able to save their work and go back and try later, means the convenience of looking up and doing things, means the easiness of navigating and locating information, means extended learning environment. Moreover, it even means it's like having your own personal tutor presenting the information to you.

Lack of Curriculum and Technical Support

The biggest challenge the students faced was lack of curriculum and technical support. Most of the participants further referred the lack of support in WBL to the fact that they couldn't ask questions and couldn't get emphatic feedback real time, which is a part of the interactivity problem. A review of the literature suggests the value of integrating interactivity into web-based learning environment, which includes questions and feedback. Weller (1988) suggested the feedback process as "the leaner actively adapts to the information presented by technology, which in turn adapts to the learner…." Northrup (2001) indicated that the reason of adding feedback as one of the interaction strategies is because of the need to "close the communications loop," the students are usually uncomfortable until they get the feedback (p. 31). This is in accordance with the findings of this study that the students need real time support for asking questions and getting feedback. This study, furthermore, adds to the literature by finding out that what the students need are not only real time feedback, but also emphatic feedback which means they want somebody to actually feel and experience what they are doing and then give them the exact feedback they need. The findings provide valuable suggestions to future design and development of WBL. Hong, Kinshuk, He, Patel & Jesshope (2001) discussed the application of Mobile agents in Web-based Learning environment which might address this need for empathic feedback.

Instructional Strategies

The students' evaluation of the instructional strategies used in WBL provides valuable suggestions for future design and delivery of WBL. The instructional strategies indicated by the students, such as leaner control and the use of step-by-step instruction, are supported by the existing literatures. Khan (2000) confirmed the value of goals and objectives and the organizations of the content in creating a meaningful learning environment. Mayer (1997) justified the effect of the use of coordinated presentation of explanation in visual format (illustrations). The findings of this research add to the literature by suggesting that the use of correspondent images is good to both novice learners and expert learners and therefore accommodates the learner differences to certain extent. One of the participants, CY1, had extensive background knowledge and experience in computers and technology, therefore, he didn't mainly use this WBL for step-by-step instructions; instead, he used it as a trouble shooting tool which helped him whenever he needed. Future design the development of the WBL instructions should consider integrating correspondent images more—especially for the more technologically advanced students.

Suggestions for Future Research

Because of the limited time available, there are several limitations of this study. First of all, this is only a preliminary research. The researcher didn't have sufficient time to do follow-up interviews with each participant. Therefore, some questions still remain unanswered. For example, CY1 mentioned sitting in front of the computer is like having "somebody" presenting it to you. But what exactly the "personal tutor" does and how to enhance "personal tutoring" in WBL are still not clear. Further follow-up interviews will be conducted to find out the answers to these questions. Secondly, the researcher only did a cursory examination of the existing literature. There is not sufficient time to interrogate the literature. A thorough examination of the literature needs to be done in the future. Thirdly, because the WBL environment in this study is mainly focused on one specific situation: the WBL instruction is a stand-alone step-by-step instruction serve as a supplement to a regular class, further research needs to be done to explore students lived experience in other kinds of WBL environments, such as an entirely online course.

References

Barron, A. (1998). Designing web-based training. *British Journal of Educational Technology, 29*(4). 355-370.

Hiltz, S.R. (1994). *The virtual classroom: Learning without limits via computer networks.* Ablex Publishing Corporation, Norwood, NJ.

Hong, H., Kinshuk, He, S., Patel, A., Jesshope, C. (2001). Application of Mobile agents in Web-based Learning environment . *Proceedings of EdMedia*, June 25-30, 778-783.

Jolliffe, A., Ritter, J. & Stevens, D. (2001). *The online learning handbook: Developing and using web-based learning.* London: Kogan Page.

Jung, I. (2001). Building a theoretical framework of web-based instruction in the context of distance education. *British Journal of Educational Technology, 32* (5)*,* 525-534.

Khan, B.H. (1997). Web-based instruction: What is it and why is it? In B.H. Khan (Ed.), *Web-based instruction*, Educational Technology Publications, Englewood Cliffs, NJ. 5-18.

Khan, B.H. (2000). A framework for open, flexible and distributed learning. *Proceedings of the Technological Education and National Development Conference, USA,* April 8-10.

LeTourneau, C. (2003). *Accessible web design*. Retrieved October 10, 2003, from: http://www.starlingWeb.com/webac.htm

Mayer, R.E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist,* 1, 1-19.

Northrup, P. (2001). A framework for designing interactivity into web-based instruction. *Educational Technology, 41*(2), 31-39.

Simich-Dudgeon, C. (1998). Developing a college web-based course: Lessons learned. *Distance Education, 19*(2), 337-357.

Van Manen, M. (1997). *Researching lived experience: human science for an action sensitive pedagogy.* (2nd edition). London, Ontario, Canada: Althouse Press.

Weller, H. G. (1988). Interactivity in microcomputer-based instruction: Its essential components and how it can be enhanced. *Educational Technology, 2* (2), 23-27.

Index