Pre-service Teachers and Search Engines: Prior Knowledge and Instructional Implications

Susan M. Colaric
Bethann Fine
William Hofmann
East Carolina University

Abstract

There is a wealth of information available on the World Wide Web that can assist pre-service teachers in their course studies. Yet observation of students in a technology integration class indicated that students were not able to find resources efficiently or reliably. The purpose of this study was to establish a baseline of what undergraduate, pre-service teachers know about search engines and searching the Web prior to beginning a course on integrating technology into the classroom. A total of 355 undergraduate, pre-service teachers over three semesters participated in the study. The results indicate low declarative, syntactic and semantic knowledge. Implications for course and program revisions based on the results are addressed.

Academic programs are continuously under evaluation and revision. Whether through a formal process from an outside agency, or the individual reflections of an instructor, these reviews seek to determine the effectiveness and value of a specific course or program in terms of the validity of objectives, relevancy and sequence of content, relation between assumed prior knowledge and current course content, and achievement of specific goals. In an internal review of a technology integration course at a southeastern university, the researchers found that faculty in the college of education assumed students at the university knew how to use the Internet as a source of information to support their academic studies. Yet the researchers observed many students had difficulty locating information relevant to their courses and assignments. The researchers also observed students spent a great deal of time on unsuccessful searchers; few students were able to locate relevant resources quickly. The purpose of this study was to establish a baseline of what undergraduate, pre-service teachers know about search engines and search strategies. The results would then be used to determine if any changes in the curriculum were warranted.

Literature Review

Many people turn to the World Wide Web as a source of information. Whether searching for information on travel, health, entertainment, or academic resources, the Web has many attractive qualities: it is easy to access, it is "open" 24 hours a day, seven days a week, there is no need to venture out in the cold or rain, and there is privacy as people search for the information they need. Yet research among the general public has shown searchers looking for information on the World Wide Web have a difficult time developing search queries and using a search engine (Chen, Houston, Sewell & Schatz, 1998; Lazonder, Biemans & Wopereis, 2000). Searches tend to be simple (Spink, Bateman & Jansen, 1999) and Boolean operators are used infrequently and incorrectly (Jansen, Spink & Saracevic, 2000). Although time spent searching the Web is high (Sullivan reports 31% of Internet users utilize a search engine two to three times per week [Sullivan, 2001]), frustration levels are also high with 71% of respondents reporting they get frustrated when searching for information on the Internet (Sullivan, 2002). Frustration and poor search skills rarely lead to positive results. It is possible, though, that college students may have more experience with computers, and thus may be more successful at web searching than the general population.

A 2001 survey among U.S. college students who use the Internet for school-related assignments found 70% of respondents reported they were successful in finding what they seek most of the time (OCLC, 2002). These students have confidence in their ability to locate information for their assignments and they find the Web easy to use. The study, however, only polled those already using the Internet. Another survey conducted in 2001, found 91 percent of students rated themselves competent in accessing information on the Internet (Osika

and Sharp, 2002). Do all college students report such success? If so, is this self-reported data true in practice? Specific to the review of this course, are pre-service teachers adept at the skills that will help them find information on the Internet?

The purpose of this study was to establish a baseline of what undergraduate, pre-service teachers know about search engines and searching the Web prior to beginning a course on integrating technology into the classroom. The following research questions were addressed:

- What factual knowledge do these students have about search engines?
- When given questions, how do these students structure a search string?
- Can these students describe how a search engine operates?
- Does knowledge of search engines vary with the amount of time a person spends on the Web?
- Does knowledge of search engines vary with the age of the student?

The study examined three types of knowledge associated with web searching: declarative knowledge, syntactic knowledge, and semantic knowledge (Colaric, 2002). Declarative knowledge (Anderson, 1982) refers to understanding facts; in this case, facts about search engines. Syntactic knowledge refers to knowledge of the language units and rules when working with a computer system (Mayer, 1992); in this case, how to structure a search query using terminology the search engine can interpret correctly. Semantic knowledge refers to the user's understanding of the major locations, objects, and actions inside a computer system (Bayman & Mayer, 1988; Mayer, 1989).

Method

Participants
Potential participants were undergraduate pre-service teachers at a research university in North Carolina. All students enrolled in a pre-service teaching course on integrating technology in the classroom during three semesters (spring 2002, fall 2003, and spring 2004) were invited to participate. The technology integration course is required for College of Education students and is usually taken during the students' junior year. Completion of the surveys was voluntary and no incentive was provided for completion.

Instrument
The survey instrument was developed previously by one of the researchers (Colaric, 2002). Eighteen questions were used to gather information in four areas: personal information, declarative knowledge of search engines, syntactic knowledge of search engines, and semantic knowledge of search engines. The questions included short answer and multiple choice. Eight questions related to participants' personal information: age, gender, number of semesters completed at the university, major field of study, minor field of study, whether he/she owns a computer, approximate number of hours per day searching the Web, and approximate number of hours per day sending/receiving email. Five questions related to declarative knowledge of search engines. Participants were asked to answer questions about whether or not all search engines work the same way, whether search engines look at all web sites on the WWW, the difference in the amount of results obtained by using AND and OR, the name of the program used by search engines to gather web sites, and whether the search term used needs to match the engine's index in order for a site to be returned. All questions included an option of "I don't know". Three questions asked participants to write down what he/she would normally type into a search engine when looking for information on a given topic (syntactic knowledge). An example is: "Suppose you want to find web sites that describe the naval battles that took place during the Napoleonic War. What would you type into a search engine?" Two questions asked the participants to describe what a search engine would do when given a particular search query. Both queries contained information grouped in parentheses, as well as use of the Boolean operators AND and OR. These two questions assessed the participants' semantic knowledge of a search engine by asking them to describe what goes on inside the system when a command is executed.

Data Collection and Analysis
The materials were administered in the pre-service teaching course on integrating technology during its normal time and day in the first or second week of classes. Surveys were coded by the researchers and results

were analyzed using SPSS (Statistical Package for the Social Sciences). Demographic data was entered as the participants recorded; groupings were also established for age (traditional undergraduate age of 19 to 22 and age 23+) and use of the Internet (both Web searching and email; "low" group of less than one hour, "medium" group of 1 to 2.5 hours, and "medium-high" group of 3 to 4.5 hours, and "high" group of 5+ hours).

Questions for the declarative knowledge section were multiple choice with answers coded and entered into SPSS. For the syntactic knowledge section, a numeric score for each answer was recorded based on three categories: accuracy of concepts identified, inclusion of variable concepts, and accuracy of Boolean expression. For the semantic knowledge section, a numeric score for each answer was recorded based on six categories: understanding of OR as a join, understanding of AND as an intersect, inclusion of all terms from the query, understanding that the search engine is querying a database set, understanding of searching for a literal string of characters, and understanding that all search engines operate in a unique manner (Colaric, 2002).

Results

Demographic Results

A total of 355 students completed the survey over the three semesters. Participants ranged in age from 19 to 57 with a mean age of 23 (median age was 21). Eighty-three percent of the participants were female; seventeen percent were male. Semesters completed ranged from one to sixteen; the mean number of semesters completed was 8.5. Ninety percent of participants own their own computers. Fifty percent of the participants were Elementary Education majors; all participants were studying in some area of teacher education. Participants reported spending an average of 1.5 hours per day searching the Web (range of 0 to 10) and 1 hour per day on email (range of 0 to 8).

Comparisons between the different semesters were tested to determine if any variations existed; the groups were found to be similar in all areas. Cross tabulations were run to determine if students of traditional college age (19-22; $\underline{n}$ = 270) spent more time on email or searching the Web than students of non-traditional age (23+; $\underline{n}$ = 84). No significant differences were found.

Declarative Knowledge

The participants in the study appeared to have some prior factual knowledge of search engines. Most ($\underline{n}$ = 355; 64%) understood search engines operate differently from each other. Of concern are the 14% ($\underline{n}$ = 49) of participants who thought search engines were all the same and 22% ($\underline{n}$ = 80) participants who did not know if search engines were all the same.

A fairly high number of participants believe search engines peruse all sites on the Web ($\underline{n}$ = 143; 40%) while a number of participants ($\underline{n}$ = 73; 21%) are not sure about this idea. Forty-seven percent of participants ($\underline{n}$ = 168) understood that terms typed into a search engine need to match the indexed sites of the engine in order to be returned. Thirty-three percent did not know ($\underline{n}$ = 115) and 20% thought this was false ($\underline{n}$ = 71). Less than half of participants ($\underline{n}$ = 162; 46%) understood the Boolean operator OR retrieves more results than the operator AND. Twenty-five percent ($\underline{n}$ = 89) thought OR retrieved less results; 23% did not know ($\underline{n}$ = 81) and 6% thought OR retrieved the same amount as AND ($\underline{n}$ = 23).

There were no statistically significant differences between the age groups (19-22; $\underline{n}$ = 270 and 23+; $\underline{n}$ = 84) in relation to declarative knowledge. Nor were there statistically significant differences in the amount of time spent searching the Web or the amount of time spent on email in relation to declarative knowledge.

Syntactic Knowledge

When asked to construct a search string given a particular topic, participants tended to construct very simple queries with a mean of 2.9 terms per query. Scores for syntactic knowledge could range from 0 to 18; the mean score was 5.36. Twenty-five percent of participants used AND correctly in their search string; 1% used AND incorrectly. One percent of participants used OR in any of the three search strings. Twenty-seven percent of participants included stop words in their queries (common words which are ignored by some search engines or result in a high number of listings for other engines). Three percent of respondents included words not directly used in the search question; in all cases the words used were appropriate synonyms for the terms in the search question. Six percent of participants used phrasing with 1% of those using it incorrectly. Truncation, NOT, and search modifiers were not used by any participants.

There were no statistically significant differences between the age groups (19-22; $\underline{n}$ = 270 and 23+; $\underline{n}$ = 84) in relation to syntactic knowledge. Nor were there statistically significant differences in the amount of time

spent searching the Web or the amount of time spent on email in relation to syntactic knowledge.

Semantic Knowledge

Participants were generally not successful in describing their semantic knowledge, scoring a group mean of .47 points out of a possible 12 (standard deviation = 0.96; range = 1, 5). Two hundred and seventy-two participants (77%) received no points for this section. Participants who did respond were slightly more likely to include a description of AND as an intersect (15%) than include OR as a join (10%). Most participants (81%) did not include all of the terms included in the question, choosing instead to describe what the engine would do with just one or two terms.

There were no statistically significant differences between the age groups (19-22; $\underline{n}$ = 270 and 23+; $\underline{n}$ = 84) in relation to semantic knowledge. Nor were there statistically significant differences in the amount of time spent searching the Web or the amount of time spent on email in relation to semantic knowledge.

Discussion

Evaluation and revision are necessary in all successful academic programs; however, the success of programs that integrate technology goals and objectives hinge on the evolution of new technological innovations as well as student's technical skill subsets. As program, accreditation and state standards are revised to correlate with the demands of the 'technology age' it is imperative that the courses that support technology integration for educators rise to meet and exceed the needs of the populations that they serve. The purpose of this study was to establish a baseline of what undergraduate, pre-service teachers know about search engines and searching the Web in an effort to evaluate the course curriculum in place and determine if changes are necessary to effectively meet the goals and objectives of the course. A current review of the literature as well as an informal review of the course indicated that students may not have the prerequisite skills needed to engage effectively in the Web searching skills necessary for the planned course content.

Confirming instructor suspicions, a significant proportion of students, 36%, surveyed indicated that they did not have the declarative knowledge necessary to effectively use search engines. These students believed that search engines did not differ in function or were not aware of their function. The results from the survey further showed that most students had misconceptions regarding Boolean operators, the construction of search strings, and were generally not successful in describing their semantic knowledge of search engines. These results corresponded with the suspicions of the course instructors noted in informal reviews and evaluations of the course but contrasted with generally held assumptions of student knowledge of web searching held by faculty in the College of Education at large. Student time on the web and demographic variations did not demonstrate a significantly statistical difference in semantic, systemic or declarative knowledge which defeats the notion that time on task correlates with efficient searching. Instead, time on the Web perpetuates poor searching. These results support the need for course revisions in curriculum programs that integrate Web technology, particularly the Educational Technology course where the survey was conducted, by establishing a baseline of student knowledge. Without an understanding of students' prior knowledge, integration of new knowledge is bound to be less successful. In addition, course revisions will focus on a curriculum alignment that works with a triangulated approach to search engines and Web searching. Units of curriculum will focus on evaluating student's prior knowledge as well as incorporating declarative, semantic and systemic approaches to instruction of search techniques as a development of baseline skills necessary to achieve the goals and objectives necessary for program, accreditation and state standards.

Sound design principles require that course developers as well as instructors evaluate the effectiveness of course materials, curriculum and objectives. In a technology based pre-service teacher course, evaluations and curriculum revisions must reflect the innovations presented in technology as well as in schools. The results from this survey suggest that curriculum revisions include professional development with in the College of Education, development of CAI tutorials for student instructional supplement, implementation of modules within the course content focusing on the online research and methods.

The use of technology in a College of Education can not be narrowed down to one specific course. Effective instruction requires that students are given the opportunity to view various methods of integration, modeling and repetition in order for retention and meaningful learning to take place. In order to facilitate and model appropriate schemas of technology integration, staff development for all College of Education Faculty will be developed and made available. The development module will contain training on Internet research techniques, Boolean operations, and technology integration methods for University faculty that correlate with ISTE's NET-S standards. These development modules will offer faculty the opportunity, time and information necessary for implementation within their course curriculums.

Development of CAI tutorials as well as the implementation of curriculum modules within the technology service course have also been recommended based on the survey results. CAI tutorials will be developed and used as course supplements for College of Education faculty as well as faculty teaching the pre-service technology course. These tutorials in conjunction with the addition of a curriculum module of web research techniques will serve as the curriculum revisions necessary to meet the needs of both the students and faculty within the pre-service technology course. Faculty will continue to re-evaluate course curriculum and student needs by developing a focused evaluation instrument designed to measure the effectiveness of instruction based on the three areas of knowledge, Declarative, Syntactic and Semantic, and their effective outcomes on meaningful learning.

A triangulated approach to further research is recommend by the researchers to determine the global impact of the relationship between the declarative, syntactic and semantic knowledge relationships effect on instruction and learning. In order to develop CAI training and curriculum modules that are instructionally effective investigation will focus on the relationships between meaningful learning and the three knowledge domains. Further research will include the development of a pre-test/post-test instrument designed to evaluate the revision changes that will be implemented in the technology course to gauge knowledge acquisition changes. These results will assist in evaluating if student needs and course objectives are being met through the implementations made through the suggestions of this research. Lateral entry teachers as well as teachers in the k-12 field will also be included in continued research efforts to determine the effects of experience and effective implementation of web searching techniques and prior knowledge.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Bayman, P. & Mayer, R. E. (1988). Using conceptual models to teach BASIC computer programming. *Journal of Educational Psychology, 80*, 291-298.

Chen, H., Houston, A. L., Sewell, R. R., Schatz, B. R. (1998). Internet browsing and searching: User evaluations of category map and concept space techniques. *Journal of the American Society for Information Science, 49*, 582-603.

Colaric, S.M. (2003). Instruction for web searching: An empirical study. *College and Research Libraries, 64*, 111-123.

Jansen, B. J., Spink, A., Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the Web. *Information Processing and Management, 36*, 207-227.

Jones, S. & Madden, M. (2002). The Internet goes to college: How students are living in the future with today's technology. Pew Internet and American Life Project, September 15, 2002. Available online: http://www.pewinternet.org/pdfs/PIP_College_Report.pdf. Accessed August 2, 2004.

Lazonder, A. W., Biemans, J. A. & Wopereis, I. G. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science, 51*, 576-581.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*, 43-6.

Mayer, R. E. (1992). *Thinking, Problem Solving, Cognition.* (2nd ed.). New York: W. H. Freeman and Company.

OCLC (2002). OCLC white paper on the information habits of college students: How academic librarians can influence students' web-based information choices. OCLC Online Computer Library Center, Inc. June, 2002. Available online at: http://www5.oclc.org/downloads/community/informationhabits.pdf. Accessed August 2, 2004.

Osika, R.E. & Sharp, D. P. (2002). Minimum technical competencies for distance learning students. *Journal of Research on Technology in Education, 34*.

Spink, A., Bateman, J. & Jansen, B. J. (1999). Searching the Web: A survey of EXCITE users. *Internet Research: Electronic Networking Applications and Policy, 9*, 117-128.

Sullivan, D. (ed.) (2001). WebTop search page study. The Search Engine Report, February 2, 2001. Online: http://searchenginewatch.com/sereport/article.php/2163451. Accessed August 5, 2004.