DOCUMENT RESUME

ED 352 949 IR 015 894

AUTHOR	Pryor, Brandt W.
TITLE	Streams of Knowledge, Streams of Action: The River of Research and Development at Technology Based Learning and Research, Arizona State University.
PUB DATE	Aug 92
NOTE	10p.; Paper presented at the Meeting of the Association of Management (Las Vegas, NV, August 4-7, 1992).
PUB TYPE	Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Descriptive (141) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Computer Assisted Instruction; Conventional Instruction; Curriculum Research; *Educational Research; Formative Evaluation; Higher Education; *Instructional Effectiveness; *Interactive Video; *Intermode Differences; Mathematics Education; Multimedia Instruction; Needs Assessment; *Pilot Projects; Preservice Teacher Education; Summative Evaluation; Videodisks
IDENTIFIERS	Arizona State University; *Technology Based Instruction

ABSTRACT

This paper describes the research agenda at Technology Based Learning and Research (TBLR) at Arizona State University, whose primary research mission is to provide information to guide the development and subsequent application of its technology based instruction (TBI) products. It is noted that the research question at the heart of the agenda is the investigation of TBI effectiveness, especially relative to conventional instruction. Each of the stages of the agenda is described--i.e., needs assessment; logic modelling; five stages of formative evaluation (monitoring, systematic review using formal criteria, presentation features analysis, open-ended expert review, and informal field testing); beta testing of the product; and summative evaluation (formal field tests and controlled tests). The various stages in the agenda are illustrated by references to a recently completed pilot project, Teaching Mathematics Methods Using Interactive Videodisc (TMMUIV), a prototype interactive, multimedia instructional tool with both classroom presentation and stand-alone laboratory applications that was developed to help preservice teachers integrate mathematics content with the teaching methods and manipulative materials advocated by the professional teaching standards of the National Council of Teachers of Mathematics. (Contains 26 references.) (ALF)

Streams of Knowledge, Streams of Action: The River of Research and Development at Technology Based Learning and Research, Arizona State University

Brandt W. Pryor
Senior Research Associate
Technology Based Learning and Research
Arizona State University
Tempe, AZ 85287-0908

Paper presented at the meeting of the
Association of Management
August 4-7, 1992
Las Vegas, NV

AUTHOR NOTE (9/1992): This version of the paper includes the concept of logic modeling which was included in the oral presentation, but not the version of the paper printed in the proceedings of the meeting.

Introduction

The ever increasing number of computers and other instructional technologies in schools, colleges, and universities today bears witness to the belief that technology can be an effective, and even revolutionary, aid to education. Is this view justified? In a recent collection of essays, entitled Imperial Masquerade, Lewis Lapham (1990) observes that:

> the belief in the sovereign powers of various new technologies has taken so firm a hold of the public imagination that it has become the stuff of primitive religion. Let the school administrator announce that he has ordered computers for eight hundred illiterate sophomores, and lo, they have become educated.

Primitive or not, this belief has been the subject of considerable investigation. Generally, research that has compared technology based instruction (TBI) with conventional instruction has found no significant differences in effectiveness between the two (e.g., Clark, 1984; Dence, 1980; Leiblum, 1982). Other studies (e.g., Atkinson, 1984; Bangert-Drowns, Kulik, & Kulik, 1985; Fletcher, 1990; Fletcher, Hawley, & Piele, 1990) have found evidence that TBI is more effective than conventional instruction.

How do we explain these discrepant results? Some scholars (e.g., Bright, 1983; Clark, 1984; Hannafin & Peck, 1988) suggest that the discrepancy might well be due to the greater care given the design of the TBI instruction that is used as comparison in these studies. Another possibility is that more recently-developed TBI is more effective than the older versions (Bangert-Drowns, et al., 1985).

It seems likely that some forms of TBI will be effective in some ways, with some learners, and there is support for the differential effects of TBI on various kinds of learners (e.g., Dalton, Hannafin, & Hooper, 1989; Falk & Carlson, 1991). If this is true--as I suspect it is--this suggests a rather intensive research effort for those of us involved in research on, and development of, TBI. Although there is a wide variety of research emphases, one of the most important is the investigation of TBI effectiveness, especially relative to conventional instruction. This research question is at the heart of our evolving research agenda at Technology Based Learning and Research (TBLR), Arizona State University. I would like to describe that agenda--in its current manifestation--and use the findings of studies of one of our recent projects to illustrate the various stages of that agenda.

Research Agenda At Technology Based Learning and Research

As a research and development unit, our primary research mission is to provide information to guide the development and subsequent application of our TBI products. Our agenda includes needs assessment, logic modelling, five stages of formative evaluation, beta testing of the product, summative evaluation, and related studies concerning applications of TBI. The various stages in the agenda will be illustrated by references to the research done on a recently completed pilot project, Teaching Mathematics Methods Using Interactive Videodisc (TMMUIV).

The TMMUIV system is a prototype interactive, multimedia instructional tool with both classroom presentation and standalone laboratory applications. It was developed to help preservice teachers integrate mathematics content with the teaching methods and manipulative materials advocated by the professional teaching standards of the National Council of Teachers of Mathematics (NCTM). The prototype system contains a database of video segments portraying exemplary elementary mathematics instruction with two manipulative devices: Geoboards and Base 10 Blocks. The TMMUIV classroom presentation mode offers both video enhancement of traditional lecture and prepared lessons. The standalone laboratory mode offers both prepared lessons and multiple opportunities for individual or small-group exploration of a large number of exemplary, elementary mathematics instructional segments.

Needs Assessment

Needs assessment can take many forms, from a literature review performed for the initial design of a new product, to field-based investigations of the learning preferences and needs of the audience for whom a given product is intended. Effective needs assessment can often guide development away from blind alleys that would otherwise be revealed only later during formative evaluation, and thus can save considerable time and other resources. Needs assessment on the TMMUIV project included reviews of the literature on teacher preparation and mathematics manipulatives, as well as investigations of the latest technology appropriate to the intended system. Toward the end of needs assessment, the next stage of our research agenda, logic modeling, will be initiated.

Logic modeling is a recent addition to our agenda, and was not performed during the TMMUIV project. We are now testing the concept's utility in two current projects. The concept is discussed separately because we now see it as a process that continues from the end of needs assessment to the beginning of summative evaluation, and perhaps even to applications research.

Logic Modeling

Origin. A major problem in project evaluation is the attempt to summatively evaluate projects that lack sufficient developmental maturity to have significant effects (Cronbach, 1987). A means of assessing the likelihood of a project having significant effects, was developed by Joseph Wholey and his colleagues at the Urban Institute (Nay & Kay, 1982; Wholey, 1979, 1987) and in the U.S. Department of Health Education and Welfare where it was used extensively. The concept was later carried to the Department of Agriculture where it has seen extensive use in planning and evaluating Cooperative Extension programs (e.g., Mayeske, 1991, 1992; Smith, 1989). It seems likely that the concept could be useful in the planning and evaluation of TBI.

Definition. Wholey's (1979) logic model (referred to by others as a program model or theory of program action: Mayeske, 1991, 1992; Smith, 1989) is a schematic diagram illustrating the events that must take place for the project to have its intended effects. This description of the structural and operational characteristics of a program's essential, manipulable components and activities (connected in a causal sequence to end effects) should be based on empirical knowledge of the content area. (Therefore, we see it beginning toward the end of needs assessment.) It provides a testable rationale for believing that certain outcomes will occur as a result of program activities. The result of this modeling is an assessment of the evaluability of a project; whether or not the intended effects are sufficiently probable for summative evaluation to be useful. (Emphasizing the result of logic modeling, Mayeske, Smith, and others refer to the process as evaluability assessment. Wholey's usage is retained in this paper.)

Purposes. The process of logic modeling could contribute both to evaluation research and development in educational technology in two ways. First, it makes explicit the developers' assumptions underlying decisions about how the project will work (beliefs about cause-effect relationships). Second, once these assumptions are explicit, any gaps in the model become apparent and can more easily be filled. Third, the process makes possible a reality-test of these underlying assumptions. The process might be useful, therefore, at many stages of our research agenda, from project monitoring, through informal field testing, and summative evaluation, perhaps even to applications research.

Development. Four steps appear necessary to develop a logic model for an educational technology:

1. Identify key components and activities
2. Determine causal sequence
3. Identify performance indicators and deadlines, and
4. Diagram a series of "if-then" statements about key components.

Only application and testing of this process will confirm its utility for research and development of TBI, but the potential appears sufficiently great to warrant this effort.

Formative Evaluation

Flagg (1990), and others have reminded us that rigorous formative evaluation is essential for effective and efficient product development. At TBLR, the information provided by formative evaluation is the primary means of assessing our development work and making needed revisions at the earliest possible moment. There are several stages of formative evaluation that we find useful. Each of our products is unique, and not all require each of the stages mentioned below, but generally we use many of the following, and generally in pretty much the same order as they are given here.

Monitoring. Monitoring of the production schedule provides useful information for improving our efficiency and effectiveness. This initial phase of evaluation is also helpful for the planning of future projects. As TMMUIV was an application of cutting edge technology, it was no surprise that our monitoring revealed the majority of delays on this project were due to later- than-expected delivery of newly developed software and hardware. Sometimes, current software did not meet our needs and had to be developed from scratch.

Systematic review using formal criteria. These reviews are performed at each appropriate stage of development, initially by TBLR staff, first members of the development team, then members of the research team. These efforts include "proofing and editing" of the presentation and de-bugging the programming. Later, outside experts in relevant fields provide another review. Sometimes the criteria are internal (e.g., developed out of the product's instructional design) and sometimes external (e.g., the professional teaching standards of the NCTM.

Presentation features analysis. This is the first stage at which data are collected from representatives of the intended audience for the product. The emphasis is on uncovering any limitations in presentation features or programming that were not discovered in-house or by our outside experts, during the previous systematic review. Time and budget restrictions prohibited the presentation features analysis on TMMUIV being performed separately from informal field testing, which is discussed below.

Open-ended expert review. A final review of the product is sometimes performed by other outside experts before we "let it out the door." Sometimes this open-ended review suggests informal hypotheses about the product's effects on users. For example, open ended review of the prototype TMMUIV system produced comments such as the following (emphases added):

> "Use of videodisc makes instruction very real "authentic" for both preservice and inservice teachers."
>
> "interactive approach permits student teachers to experience classroom situations. Preservice experience can modify the actual classroom response and management."

Informal field testing. Once we have developed the product as fully as possible, given the limits of our knowledge, we expand that knowledge by testing the product with conditions and learners as similar as possible to those for which it is finally intended. This "acid test" often provides much useful information. Sometimes, problems that were undetected in previous stages of formative evaluation appear only when the product is tested under field conditions. Sometimes the resulting data lead us to informal hypotheses about the product's effects on learners and how those are achieved, or provide confirmatory or disconfirmatory evidence concerning hypotheses developed at previous stages of evaluation. For example, interviews with preservice teachers following the informal field tests of TMMUIV suggested that preservice teachers gained a "surrogate for experience" in teaching with the Geoboard, echoing the above comments of expert reviewers. For example (emphases added):

> "Rather than just learning theory, we saw the material actually being applied in the classroom."
>
> "It's sort of like getting experience. How do you teach experience?"
>
> It makes you a bit more qualified than other people already teaching who don't have experience with them [Geoboards]."

When the system was sufficiently developed to permit an exploratory summative evaluation, this variable of "surrogate experience" was refined for testing in subsequent stages of the evaluation.

Beta-Testing

In all previous stages of evaluation, the product is used either by the development team, in their presence, or by persons trained by them. Beta-testing provides a yet more rigorous evaluation of the product: Can the "naive user" operate the system effectively? The information gleaned from this activity can be most useful in making final revisions of the product. TMMUIV has not yet been formally beta-tested , but we have approximated this stage of evaluation by having "naive" users operate the system without special training. The difficulties they encountered helped us make useful revisions in our without special training. The difficulties they encountered helped us make useful revisions in our TMMUIV User's Guide.

Sometimes the results of beta-testing lead to important design changes in the product. TMMUIV was originally conceived as a database for classroom presentation or independent viewing of exemplary instructional video segments. We found that many mathematics methods instructors at Arizona State University (our "naive users") were unable to devote the time necessary to develop the new instructional strategies needed to use the system effectively. This finding prompted us to design "model lessons" that could be used immediately by any instructor or by preservice teachers in the laboratory. Before the prototype is developed into a completed system, an "authoring module" will be developed to facilitate creation of custom tailored lessons by methods instructors. Formal beta-testing at other universities is expected to provide useful information for final revision of the system as well as data on product effects for designing the final summative evaluation and related studies.

Summative Evaluation

Summative evaluation is intended to provide us with information about the effects of the product, especially how these differ from those of conventional instruction. Such information is important to potential commercial distributors and end-users. Although summative work at TBLR can seldom be shoehorned into one category, we see two "ideal types" of summative evaluations: Formal field tests with more rigor than our informal field testing, and controlled tests.

Formal field tests. The knowledge gained through informal field tests and beta-testing, provides the basis for what we call formal field tests. These test the product under real conditions and administer formal measures of the supposed effects of the product. This initial stage of summative evaluation is a "confirmatory" study that attempts to find evidence supporting the "hunches" produced in the earlier work.

Controlled tests. Controlled tests have at least informal hypotheses and attempt disconfirmation of these using experimental or quasi-experimental design. By fall 1991, the TMMUIV system was sufficiently developed to permit an exploratory summative evaluation with a quasi-experimental design that is somewhere between these two ideal types. That research will be discussed below in the final section of this paper.

Applications Research

Once summative evaluation has isolated the product's effects on users, further research can be conducted to determine the longevity of these effects, how they are influenced by other factors, differential effects on users, and related areas of research. Several applications studies are now being planned for the future, including an investigation of attitude, following procedures described by Pryor (1992, April).

Exploratory Summative Evaluation of TMMUIV

This section of the paper provides an overview of the summative evaluation conducted in the fall of 1991. A more detailed preliminary report of these results is available in Pryor (1992, March). The final report of all research findings (1990-1991) will be available in Bitter (in press).

Objective

The purpose of the study was to assess the degree of added instructional value that might be attributed to the TMMUIV classroom presentation mode as an enhancement of conventional lecture. We believed that to increase the likelihood of preservice teachers instructing effectively with NCTM standards and the Geoboard after graduation, they must learn how to do so, feel prepared to do so, and be motivated to do so. We suspected that if TMMUIV influenced these variables--and if there was a causal relationship among them--that cognitive gain must occur first, and be followed by preparedness and motivation.

The three central variables of the study were therefore (a) cognitive gain needed to teach, (b) preparedness to teach, and (c) motivation needed to teach, with Geoboards after graduation. We believed that the ability to analyze classroom instructional interactions was an important aspect of cognitive gain, and the formative stages of evaluation suggested that the system gave preservice teachers a "surrogate for experience" by enhancing their ability to analyze instructional interactions. This variable, named "observational power," became an important measure of the cognitive gain that the system might bring about.

Method

Four mathematics methods classes were assigned to treatment and control groups. Each group received two hours of identical instruction, from the same outside instructor, except that the one-hour lecture segment was enhanced by TMMUIV for the treatment group. The TMMUIV prototype system was used for less than an hour and less than 12 minutes of video material were shown. Data were collected directly after the instruction by pencil-and-paper instruments.

Cognitive gain was measured in two ways. Six "retrospective pretest-" posttest items assessed changes in knowledge and competence to teach. (These items were suggested by Campbell & Stanley, 1966; similar measures have been used by Curry & Purkis, 1986, and Pryor, 1989.) Five of the items were specific-knowledge referent (e.g., competence to convey idea of "estimation," knowledge of how to help children feel comfortable, etc.). The sixth was more general, "Knowledge of what it is like in the elementary classroom."

Another form of cognitive gain--"observational power"--was assessed by three objective performance tests (posttest only) that asked subjects to analyze instructional interactions shown in videoclips--neither group had seen any of the clips before--and record their analyses in an open-response form. These items were worded as follows:

(1) "How did the teacher indicate to the students that it was all right for them to give an answer that was something other than the 'correct' answer?"

(2) "What important teaching strategy did the teacher use to teach the concept of 'area'?"

(3) "What did the teacher do that shows a way of making young children comfortable in the classroom?"

Results and Discussion

Pretreatment differences. Available pretreatment data on the subjects in each of the four methods classes suggested that control group subjects were higher than those in the treatment group on an estimated measure of academic "ability" derived from ACT/SAT scores and GPA. (This difference only approached statistical significance.) However, we didn't know--until after data collection and analysis--that the control group also had more previous exposure to instruction with Geoboards than the treatment group.

Subjects in the control group were more likely than those in the treatment group to have seen Geoboards used in their practicum schools, and significantly more likely to have had previous training in teaching with Geoboards, and to have been taught with Geoboards as elementary students. These data suggested to us that our treatment group might be at a relative disadvantage in learning about instructing with Geoboards.

Posttreatment Differences. Nevertheless, despite what appear to be potentially important learning disadvantages, and a very brief treatment with a prototype system, the treatment group outscored the control group on every single measure of the central variables in the study: cognitive gain, preparedness to teach, and motivation to teach with Geoboards. As would be expected, given the prototype nature of the system and its brief use (with less than 12 minutes of video), between-group differences were not statistically significant on motivation or feeling of preparedness. Both groups showed highly significant gains on the pretest-posttest items, and the treatment group's gain was higher on each of the six items, but only one between-group difference on gain was significant. The mean gain, on all six items, made by the treatment group was significantly higher however, and is illustrated in Figure 1.

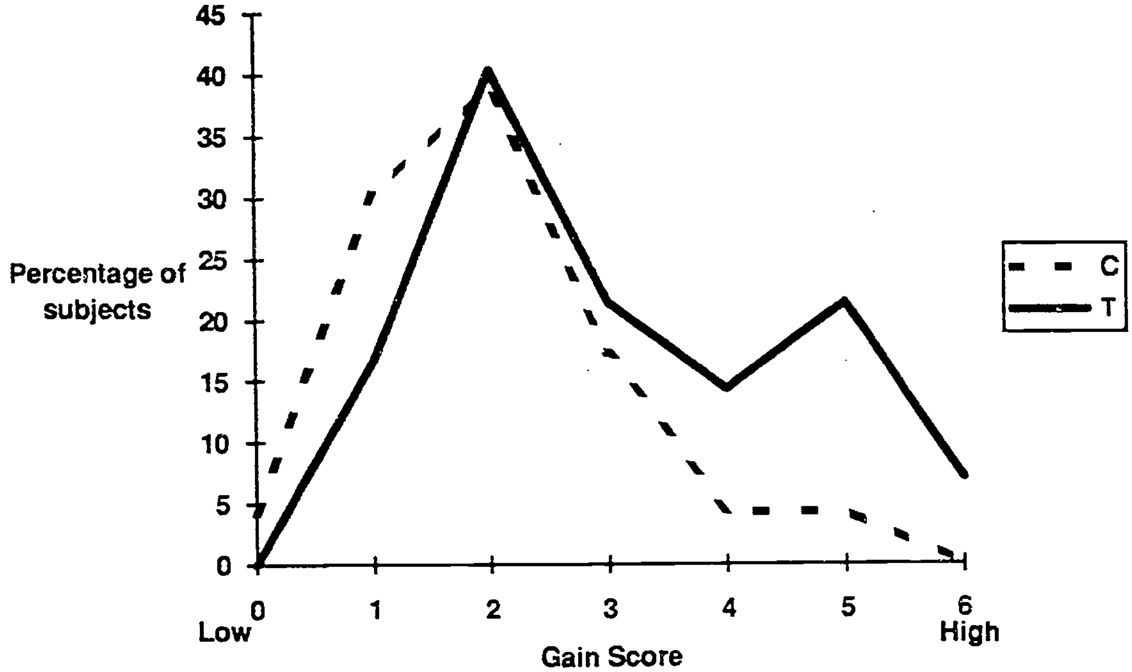

Figure 1: Mean gain made by the treatment and control groups on the pretest- posttest items.

The difference between groups on the general knowledge pretest-posttest item, "knowledge of what it is like in the elementary classroom," was significant. This difference appears related to the results on the objective performance measures. The treatment group outscored the control group on all three of the objective tests of "observational power." The overall between-group difference was highly significant and is illustrated in Figure2.

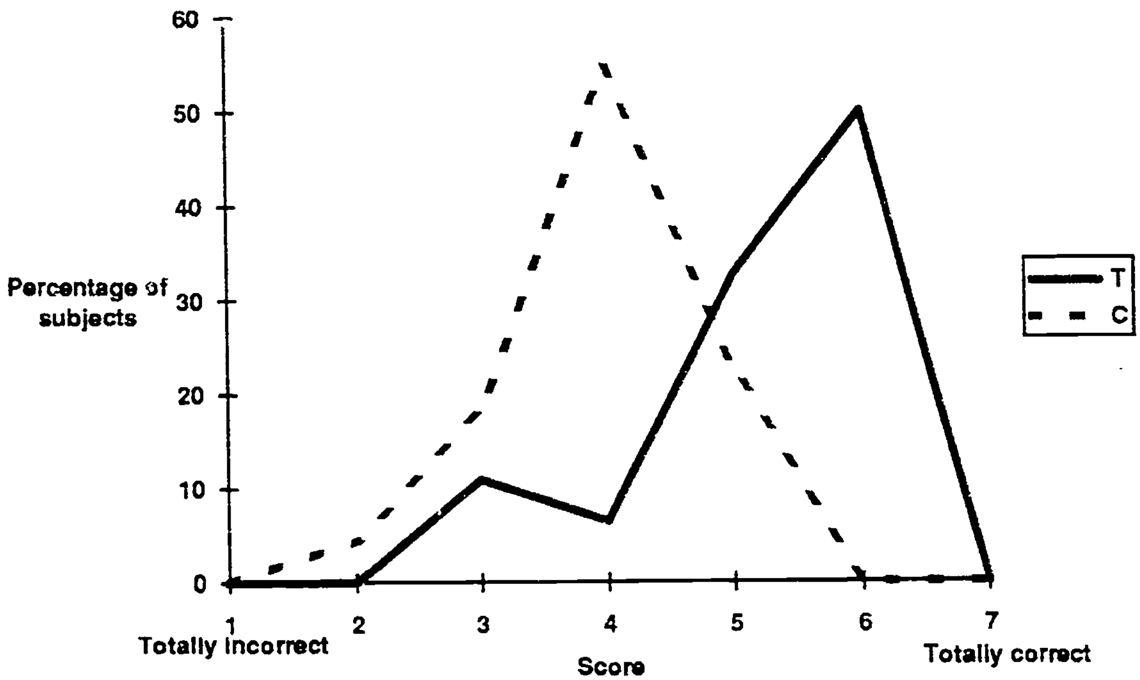

Figure 2: Mean scores of treatment and control groups on the performance tests.

6

Conclusion

This exploratory summative evaluation found considerable evidence that even brief use of the prototype TMMUIV system had significant added instructional value compared with conventional instruction alone. The subjects who were taught with it outscored arguably more able and advantaged peers on three objective performance tests. The earlier formative evaluations suggested that the completed system could have potential for inservice teachers as well, and this possibility will be pursued in future investigations. After seeing the system, one preservice teacher commented, regarding her future teaching:

> "I like the idea that, when I want to teach something I haven't taught before, I can just go into a lab and look at examples [of such instruction] on the computer. It would sure be easier than having to go to a class, or sign up for inservice, or something."

Experts who viewed a demonstration of the system said:

> "[Has] ability to use with preservice and inservice teachers [to] demonstrate classroom climate [and] content instruction...."

> "...makes instruction very real 'authentic' for both preservice and inservice teachers...."

The potential utility of this form of TBI for inservice teachers will be pursued in future research and development projects. The development of the prototype system into a completed instructional supplement in the next stage of development will provide numerous opportunities for research on its utility for preservice and inservice teachers' education. Comparisons of the system's effects on the two target audiences could lead to important advances in the knowledge base of teacher education as well as TBI.

References

Atkinson, M.L. (1984). Computer-assisted instruction: Current state of the art. Computers in the Schools, 1(1), 91-99.

Bangert-Drowns, R.L., Kulik, J.A., & Kulik, C.C. (1985). Effectiveness of computer-based education in secondary schools. Journal of Computer-Based Instruction, 12(3), 59-68.

Bitter, G. G. (in press). The research results on TMMUIV: Formative and summative project evaluation. (TBLR Monograph No. 5). Tempe: Arizona State University, Technology Based Learning and Research.

Bright, G. W. (1983). Explaining the efficiency of computer assisted instruction. AEDS Journal, 16(3), 144-153.

Campbell, D. T., & Stanley, J. C. (1966). Experimental and quasi-experimental designs for research. Boston, MA: Houghton Mifflin.

Clark, R .E. (1984, April). Learning from computers: Theoretical problems. Paper presented at the meeting of the American Educational Research Association, New Orleans.

Cronbach, L. J. (1987). Designing evaluations of educational and social programs. SanFrancisco, CA: Jossey-Bass.

Curry, L., & Purkis, I. E. (1986). Validity of self-reports of behavior changes by participants after a CME course. Journal of Medical Education, 61, 579-584.

Dalton, D.W., Hannefin, M. J., & Hooper, S. (1989). Effects of individual and cooperative computer-assisted instruction on student performance and attitudes. Educational Technology Research and Development, 37(2), 15-24.

Dence, M. (1980). Toward defining the role of CAI: A review. Educational Technology, 20(11), 50-54.

Falk, D. R., & Carlson, H. L. (1991). Evaluating the effectiveness of multimedia applications in human service and teacher education. Multimedia Review, 2(3), 12-17.

Flagg, B. N. (1990). Formative evaluation for educational technologies.Hillsdale, NJ: Lawrence Erlbaum.

Fletcher, J. D. (1990). Effectiveness and cost of interactive videodisc instruction in defense training and education. Alexandria, VA: Institute for Defense Analysis.

Fletcher, J. D., Hawley, D. E., & Piele, P. K. (1990). Costs, effects, and utility of microcomputer assisted instruction in the classroom. American Educational Research Journal, 27, 783-806.

Hannafin, M. J., & Peck, K. L. (1988). The design, development, and evaluation of instructional software. New York: Macmillan.

Lapham, L. (1990). Imperial masquerade. New York: Grove Weidenfeld.

Leiblum, M. D. (1982). Factors sometimes overlooked and underestimated in the selection and success of CAL as an instructional medium. AEDS Journal, 15(2), 67-79.

Pryor, B. W. (1989). Three longitudinal impact evaluations: Cognitive gain, performance gain, and participant satisfaction. Unpublished manuscript.

Pryor, B. W. (1992, March). Does microcomputer based instruction make a difference?: Evaluating gain in knowledge, performance, and motivation. Paper presented at the Twelfth Annual Microcomputers in Education Conference, Tempe, AZ.

Pryor, B. W. (1992, April). Measuring and understanding attitude: A more nearly complete measure for survey research. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Mayeske, G. W. (1991). Program design: An evaluation approach (Parts I and II). Washington, DC: Planning, Development, and Evaluation Staff, Extension Service, U.S. Department of Agriculture.

Mayeske, G. W. (1992). Evaluability assessments of five rural economic development programs: A synthesis. Washington, DC: Planning, Development, and Evaluation Staff, Extension Service, U.S. Department of Agriculture.

Nay, J. N., & Kay, P. (1982). Government oversight and evaluability assessment. Lexington, MA: Lexington Books.

Smith, M. F. (1989). Evaluability assessment: A practical approach. Boston, MA: Kluwer Academic Publishers.

Wholey, J. S. (1979). Evaluation: Promise and performance. Washington, DC: Urban Institute.

Wholey, J. S. (1987). Evaluability assessment: Developing program theory. In L. Bickman (Ed.) Using program theory in evaluation, New Directions for Program Evaluation, No. 33. San Francisco: Jossey-Bass.