DOCUMENT RESUME

ED 355 914 IR 015 992

AUTHOR Fletcher, J. D.
TITLE Cost-Effectiveness of Interactive Courseware.
INSTITUTION Institute for Defense Analyses, Alexandria, VA.
REPORT NO UTP-2
PUB DATE Dec 92
NOTE 41p.
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adult Education; *Computer Assisted Instruction; Computer Software Evaluation; *Cost Effectiveness; *Courseware; *Effect Size; Elementary Secondary Education; Higher Education; *Instructional Effectiveness; *Interactive Video; Learner Controlled Instruction; Meta Analysis; Military Training; Optical Data Disks; Tables (Data); Vocational Education
IDENTIFIERS *Interactive Computer Systems; Interactive Videodisks

ABSTRACT

What is known about the cost effectiveness of interactive courseware (ICW) is reviewed, and issues that remain are summarized. Effect size is used for reporting the effectiveness of ICW programs. Two ICW media are considered: computer-based instruction and interactive videodisc instruction. Effect sizes for computer-based instruction have been found to be 0.47 standard deviation (SD) over 28 evaluation studies performed in elementary schools, 0.40 SD for 42 studies in secondary schools, 0.26 SD for 101 studies in colleges, and 0.42 SD over 24 studies in adult education settings. Effect sizes for interactive videodiscs have been found to be 0.69 over 14 studies in colleges, 0.51 over 9 studies in industrial training, and 0.39 over 24 studies in military training. Across many instructional settings and subject matters, ICW programs have been found to be more effective than conventional instructional programs. Favorable cost ratios have been found for ICW programs across many settings and subject matters, but data for a conclusive case for ICW programs remain limited and incomplete. There are promising results for ICW costs and effectiveness data considered separately, but conclusive evidence awaits more complete study. Five tables report study findings. (Contains 43 references.) (SLD)

Cost-Effectiveness of Interactive Courseware

J. D. Fletcher
Institute for Defense Analyses
Alexandria, Virginia 22311

December 1992

The Technical Cooperation Program

Subgroup U

UTP-2

ACKNOWLEDGMENTS

The author gratefully acknowledges the assistance and review provided by Peter Clark (Australia), Lochlan Magee (Canada), Athol Forrest (New Zealand), Keith Ellis (United Kingdom), and Wayne Waag (United States) in the preparation of this paper.

Cost-Effectiveness of Interactive Courseware

Summary

I. Background

As managers, developers, and users of interactive courseware (ICW) become familiar with the medium, they increasingly want to know if the costs of ICW repay its users with commensurably increased instructional effectiveness. In short, is ICW a cost-effective alternative for instruction? This paper briefly reviews what we know about the cost-effectiveness of ICW programs and the issues that remain.

II. Economic Analysis

Economic analysis is a technique to help allocate resources efficiently (Okun, 1970). An assumption in the economic analysis of instructional programs is that there is a combination of resources that maximizes instructional productivity with minimum resources. Maximum instructional efficiency is said to be attained when this combination of maximized productivity and minimized resources has been achieved.

Economic analysis incorporates the concepts of both cost-benefit and cost-effectiveness analysis. The assumption underlying cost-benefit analysis is that both costs and benefits are measurable in the same units, which are usually monetary units such as dollars. Many instructional benefits are not well assessed in monetary units. When these are lacking, cost-effectiveness analysis may be used. Benefits in cost-effectiveness analyses are usually measured in their own, non-monetary units.

III. Effectiveness of Interactive Courseware

A. Effectiveness Measurement

Instructional effectiveness usually concerns whether or not a given program of instruction does what it is supposed to do -- whether it accomplishes its objectives. The objectives are assumed to be valid as given.

There are usually several objectives that a program of instruction is supposed to accomplish. In practice, few and often single measures of merit are used to determine the effectiveness of an instructional program. However, if we combine the results of a number of studies that assess the effectiveness of instructional approaches such as those available for ICW programs, the picture of their effectiveness that emerges may be reasonably accurate and sufficient for most purposes.

A method, called meta-analysis, for quantitatively combining the results of many studies has been developed over the last 15 years. Meta-analysis uses a measure called effect size, which is defined as the difference between the means of two treatment groups divided by either the control group standard deviation or a pooled control group and experimental group population. Roughly, an effect size of 0.50 suggests an improvement in student performance from the 50th to the 69th percentile. Effect size is used in this paper for reporting the effectiveness of ICW programs.

A. Effectiveness Results

Two ICW media are consider in this paper: computer-based instruction and interactive videodisc instruction. Effect sizes for computer-based instruction have been found to be: 0.47 standard deviations over 28 evaluation studies performed in elementary schools, 0.40 standard deviations over 42 studies performed in secondary school, 0.26 standard deviations over 101 studies performed in colleges, and 0.42 standard deviations over 24 studies performed in adult education settings. Effect sizes for interactive videodisc instruction have been found to be 0.69 over 14 studies performed in colleges, 0.51 over 9 studies performed in industrial training, and 0.39 over 24 studies performed in military training. The lower effect size for military training may be partially due to greater variance in military populations. There were six studies of interactive videodisc instruction in which different levels of interactivity were compared. In all six studies, effect sizes were higher for the more interactive approach.

Across many instructional settings and subject matters, ICW programs have been found to be more effective than conventional instructional programs.

IV. Costs of Interactive Courseware

A. Cost Measurement

Costs reported for any approach will vary widely depending on the assumptions and procedures used to collect -- or ignore -- different costs. What is required in measuring the costs of instruction is a practicable list, or structure, of well defined cost components. These have been provided in two cost models, the most complete of which was devised by Knapp and Orlansky (1983), who defined 75 cost elements for three major categories of cost: Research and Development, Initial Investment, and Operating and Support. This model appears to be both comprehensive and practicable for most training applications in the military.

B. Cost Results

The cost results reported by Orlansky and String (1979) still appear to be the best available for computer-based instruction. Although they found cost data to be limited and incomplete, they concluded that the cost of computer-based instruction was less than the cost of conventional instruction due to an average 30 percent savings in student time. Fletcher (1990) determined cost ratios (ratios of interactive videodisc instruction costs over conventional instruction costs) for 13 interactive videodisc instruction programs. All 13 ratios were under 1.00 indicating lower costs for interactive videodisc instruction in each measured instance. Five of these cost ratios were for initial investment costs and averaged 0.43; eight were for operating and support costs and averaged 0.16.

Favorable cost ratios have been found for ICW programs across many settings and subject matters. However, the data required for a conclusive case in favor of ICW programs remain limited and incomplete.

V. Cost-Effectiveness of Interactive Courseware

A. Cost-Effectiveness Measurement

The main issue in cost-effectiveness analyses of instructional programs appears to be that both cost and effectiveness data must be collected under the same empirical design using systematic models of cost inputs and effectiveness outputs. Two reasons for this are, first, that the same experimental controls should hold for all conditions under which the cost and effectiveness data will be drawn, and, second, that the same cost models should be used for all conditions and related to effectiveness measures in the same way across all conditions. Kazanowski (1968) developed a standardized, 10-step approach to cost-effectiveness evaluations that can be applied to interactive courseware assessments.

B. Cost-Effectiveness Results

The Orlansky and String (1979) study discussed earlier appears to remain the primary source of information on the cost-effectiveness of computer-based instruction in military training. They reported that (a) student performance under CBI was at least as good as it was under conventional instruction and (b) students reached the desired levels of performance in about 30 percent less time using computer-based instruction. A 33 percent time savings in favor of computer-based instruction was reported by Spindler (1990), and a 30 percent time savings was also reported in the interactive videodisc instruction review. Because the costs for interactive videodisc instruction were generally lower than those reported for conventional instruction in the studies reviewed and because interactive videodisc instruction effect sizes were generally higher than those reported for conventional instruction, a "suggestive" finding that interactive videodisc instruction is more cost-effective than conventional instruction was reported.

Cost-effectiveness is a relative measure and any assessment of the costs of ICW programs must list what alternatives to ICW are being considered. Only one study was found in which cost-effectiveness ratios were calculated and reported for a number of instructional approaches including peer tutoring, non-peer tutoring, reducing class size, increasing instructional time, and providing computer-based instruction for 10 minutes per day. Computer-based instruction based on micro-computers placed in classrooms was found to have the lowest cost-effectiveness ratio. The study concerned elementary school mathematics achievement, and its implications for military training are promising, but not conclusive.

There are promising results on both ICW costs and effectiveness considered separately, but conclusive evidence awaits more complete study in which both cost and effectiveness data are drawn from the same study.

VI. Utility Analysis

Instructional outcomes may not be equal in the view of decision makers. Choices among instructional alternatives may be guided by the weights decision makers assign to different instructional outcomes. Utility analysis takes account of these values. It is conceptually and procedurally simple, and consideration should be given to its inclusion in cost-effectiveness analyses.

VII. Remaining Issues

The implications of evaluations performed by ICW developers, uniqueness of ICW materials, redesign of materials for ICW presentation, production quality, relation of design to outcomes, and linkage to job performance need to be better understood in the current state of knowledge.

The adoption of a comprehensive cost model, development of baseline measurement, and establishment of a database of instructional costs are issues that remain to be resolved in the assessment of ICW costs.

The development of a media selection database, continuing need for cost-effectiveness studies and data, application of cost-effectiveness to the full system required to produce instruction, development of a standard measure for combining effect size with costs, and inclusion of cost-utility assessment are issues that remain to be resolved in the assessment of ICW cost-effectiveness.

VIII. General Recommendations

Four recommendations for cost-effectiveness analyses in military training have appeared frequently during the last 5-10 years. As articulated by Orlansky (1992) they are: (1) Include trade-offs between costs and effectiveness in cost-effectiveness evaluations; (2) Include factors of learning and forgetting in cost-effectiveness evaluations; (3) Develop databases on costs, effectiveness, and cost-effectiveness of training systems; and (4) Give high priority to investigations of military effectiveness. The recommendations remain valid and deserving of attention.

Cost-Effectiveness of Interactive Courseware

I. Background

As managers, developers, and users of interactive courseware (ICW) become familiar with the medium, they increasingly want to know if the costs of ICW repay its users with commensurably increased instructional effectiveness. In short, is ICW a cost-effective alternative for instruction? This is an important question for those who are responsible for the proper allocation of scarce instructional resources and for those who are concerned that these resources provide as much Defense capability and effectiveness as possible. Answers to this question remain incomplete, and issues associated with the question remain unresolved. This paper briefly reviews what we know about the cost-effectiveness of ICW programs and the issues that remain.

Interactive courseware is a relatively new term intended to encompass all interactive approaches to instruction. It includes computer-based instruction, interactive videodisc instruction, instruction using CD-ROM, and instructional simulation. The key distinction made between an ICW program and other training programs is the provision of interactions that tailor the instruction to the needs of individual students. With individually tailored instruction, each student receives the level of detail, pace, remediation, sequence of topics, and interactions needed to learn the material efficiently within constraints imposed by time and access to instructional resources. Notably, ICW programs can individualize instruction within our current, group-oriented instructional settings.

Since the introduction of ICW programs in the late 1950s, considerable data have been gathered on their effectiveness. Cost data have been accumulating more slowly, but some are now available. Analyses of ICW cost-effectiveness may now be possible.

The foundation for cost-effectiveness analysis is economic analysis, which is discussed briefly in Section II. Effectiveness is discussed in Section III, which reviews issues in measuring the effectiveness of ICW programs and what we have learned by measuring it. Cost is discussed in Section IV in a similar manner, beginning with issues of measuring costs of ICW programs followed by a brief review of what we have learned. Section V discusses cost-effectiveness in a similar manner. Cost-utility analysis may be a necessary step beyond cost-effectiveness and it is discussed in Section VI. Section VII lists issues that remain in assessing the effectiveness, costs, and cost-effectiveness of ICW programs used in military training. Section VIII concludes with some general recommendations for cost-effectiveness analyses in military training.

II. Economic Analysis

Economic analysis is a technique to help allocate resources efficiently (Okun, 1970). An assumption in the economic analysis of instructional programs is that there is a combination of resources that maximizes instructional productivity with minimum resources. Maximum instructional efficiency is said to be attained when this combination of maximized productivity and minimized resources has been achieved. Obviously, maximum instructional efficiency is more frequently desired than achieved.

Economic analysis incorporates the concepts of both cost-benefit and cost-effectiveness analysis. Cost-benefit analysis is generally used to determine if the benefits of projects and policies outweigh their costs. The assumption underlying cost-benefit analysis is that both costs and benefits are measurable in the same units, which are usually monetary units such as dollars. This commensurability is a prerequisite for cost-benefit analysis.

Many instructional benefits are not well assessed in monetary units, despite heroic attempts by economists and instructional evaluators to do so. When commensurability is lacking, cost-effectiveness analysis is used. In both cases, costs are measured in monetary units. However, benefits in cost-effectiveness analyses, such as information retention, the productivity and motivation of workers, supervisor ratings, and the productivity and effectiveness of the client organization, are usually measured in their own, non-monetary units.

In performing cost-effectiveness analyses for instruction a common practice is to hold either costs or effectiveness constant and observe variations in the other variable across the alternatives being considered. Often the variable is not actually held constant, it is simply assumed to be the same across all alternatives. Frequently, no data are presented to support the assumptions of equal costs or equal effectiveness, and decision makers must take it on faith that the assumptions are warranted.

Evaluations are of more benefit to decision makers when they consider both costs and effectiveness data. However, there are issues in both procedures and practice to be resolved. On the cost side, no standardized methodology for analysis of instructional costs has been widely adopted, nor are cost data routinely acquired in accord with a commonly held set of definitions. On the evaluation side, the objectives of many instructional programs are poorly articulated, not measured, or not measurable, leaving assessment without a foundation. In training, which is done to effect job performance, instructional effectiveness measures are frequently based on end of instruction test scores rather than job performance measures. Some of these shortcomings are unavoidable, some are not. Those that are not deserve some attention.

III. The Effectiveness of Interactive Courseware

A. Issues in Measuring Effectiveness

Assessment of any innovation is naturally concerned with effectiveness -- can the innovation accomplish what we expected? Effectiveness by itself is a reasonable issue for research and early development of instructional approaches. Instructional decision makers must begin with the separate data on costs and effectiveness that are available from researchers and developers before making hard decisions about cost-effectiveness.

Instructional effectiveness usually concerns whether or not a given program of instruction accomplishes its objectives. Effectiveness in this sense can be contrasted with efficiency, which concerns whether or not a given program of instruction does what it is supposed to do with minimal cost. In this section, we are concerned with effectiveness alone.

We may also distinguish between instructional validity and instructional effectiveness. Instructional objectives that reflect the knowledge and skills required by the job students are being trained to perform are said to be valid. Although problems of validity are common in training, in this section objectives are assumed to be valid as given. If an instructional program accomplishes its objectives, it is considered to be effective.

Instructional objectives may include speed of response, accuracy of response, short and long term retention of both performance and knowledge, ability to transfer performance and knowledge to new situations, insight and the ability to teach others what was learned, adherence to procedure, and motivation to pursue development of performance and knowledge in the subject area. These objectives must compete with one another for limited resources and time within any program of instruction. How much weight to assign to any one of them depends on the intentions of decision makers. Different measures might be taken and weighted differently to evaluate the same instructional program depending on these intentions.

In practice, few and often single measures of merit are used to determine the effectiveness of an instructional program. The measures may not reflect the objectives of the instructional program, and their relation to job performance may be too distant to determine the validity of the program. On the other hand, the measures commonly found in instructional evaluation studies are multidimensional and may reflect a number of instructional objectives; they are often collected with care and systematic attention to the demands of empirical methodology; and they may possess at least face validity as measures of job performance. If we combine the results of a number of studies that assess the effectiveness of instructional approaches, such as those available for ICW programs, the picture of their effectiveness that emerges may be reasonably accurate and sufficient for most purposes.

The methodology used for analytic reviews of this sort has changed considerably in the last 15 years. The "box-score" approach which earlier characterized the methodology has been replaced by "meta-analysis." In the box-score approach studies in which an experimental group exposed to the treatment under review are collected, the proportion of studies in which the experimental group means exceed control group means by some statistically significant extent is calculated, and the treatment is reported as favorable or not depending on whether this proportion is large or small. Hedges and Olkin (1980) have shown that the box score approach has very low power (low ability to detect statistically significant differences) for the treatment effect sizes and sample sizes characteristic of instructional research. They also showed that the power of the box score approach <u>decreases</u> as the number of studies included in the review increases.

Glass (1976), among others, proposed an alternative approach. Since he was performing an analysis of analyses, he described his approach as "meta-analytic." It differs from the box-score approach in three ways: (1) studies relevant to the issue at hand are collected using clearly defined procedures that can be replicated; (2) a quantitative measure, "effect size," is used to tabulate the outcomes of all the collected studies including those with results that are not statistically significant; and (3) statistical procedures are used to synthesize the quantitative measures and describe the findings of the analysis. Glass's approach appears to be especially appropriate for synthesizing the results of instructional research, and it has been widely used for this purpose since its introduction.

Effect size is defined as the difference between the means of two groups divided either by the standard deviation of the control group or the standard deviation of the control and experimental groups pooled together (Glass, McGaw, and Smith, 1981). Hedges and Olkin (1985) showed that, for every effect size, both the bias and variance of its estimate are smaller when the standard deviation is obtained by pooling the sample variance of the experimental and control groups instead of using the control group standard deviation by itself. Effect sizes based on pooled standard deviations are used in this paper.

The position of this paper is that: (1) a reasonably accurate picture of the effectiveness of ICW programs used in military training can be found through a quantitative combination of results from existing evaluation studies; (2) meta-analysis, and particularly the measure of "effect size," should be used perform this analysis; and (3) the choice of studies included in such an analysis can be properly based on the media used to present the instruction.

Positions (1) and (2) are mildly controversial, but generally accepted. Position (3) is also generally accepted, but more controversial. ICW programs are usually bounded and defined by the interactive media -- the hardware systems -- they use. Evaluations of ICW are thereby subject to the cautions raised by many commentators and best articulated by Clark's (1983) critique of media-based research. These concerns may be summed up by the notion that hardware alone does not define an instructional approach -- what is done with the hardware is what counts. This point of view seems unequivocal. The presence of computer control in a system is no guarantee of effective instruction nor even that the unique features of the system will be used.

However, critics of media effectiveness research push the argument farther. Clark states (page 445) that, "The best current evidence is that media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition." This statement may go too far. Improvements in the technology of delivering food from centers of production to markets has had a tremendous impact on the nutrition of nations. The technologies by themselves do not guarantee this impact, but the functionalities they support and their applications do. The development of movable type did not by itself have an impact on learning, but the functionalities it supported and their application to instruction did -- no books, no mass education.

Moreover, evaluation is performed for a reason, usually because there is a decision to be made based on the evaluation. An evaluation that does not serve the decision that motivates it in the first place is flawed, regardless of the soundness of its methodology. In military training, many issues center on whether investment in an instructional medium should continue. Investments in computer-based instruction, computer-managed instruction, interactive videodisc instruction, CD-ROM based instruction, and instructional simulation have all been questioned, and they are all forms of ICW. Evaluations that do not address these decisions fail to serve decision makers, instructional practice, and the state of the art.

B. What Do We Know about the Effectiveness of Interactive Courseware?

This paper pulls together some findings and attempts to determine their implications. Two ICW media are considered: computer-based instruction and interactive videodisc instruction. They are chosen because that is where sufficient effectiveness data can be found. Results from education settings as well as training settings are included for the same reason. Also, it is assumed that training and education are opposite ends of the same dimension which may be called instruction, that training includes aspects of education, education includes aspects of training, and that differences between the two stem from job-based versus career-based objectives.

The effectiveness of three individualizing systems of instruction that are not ICW is reviewed to provide a baseline for ICW programs. These systems, along with their effect sizes and the number of studies on which the effect sizes were based, are listed in Table 1. Most of these results are taken from studies performed by James Kulik and his colleagues at the University of Michigan.

The approaches in Table 1 are print oriented. They divide instructional content into units of instruction. Then, for each unit, they provide a pre-assessment of the learner, an individualized prescription of instruction content and presentations based on the pre-assessment, the instruction itself, and post-assessment(s) on which the learner must display criterion levels of knowledge and/or performance before progressing to the next unit of instruction.

Personalized System of Instruction. Keller's Personalized System of Instruction (PSI) was initiated by his 1968 paper, "Goodbye Teacher ... ." PSI has been used primarily to replace lecture-based, classroom teaching in higher education. Keller listed five features that distinguish PSI from other instructional systems: (a) the unit mastery requirement; (b) student self-pacing; (c) student proctors; (d) reliance on written instruction; and (e) de-emphasis on lectures. PSI separates instructional content into content units that are presented in a linear sequence, and it requires students to demonstrate mastery of each unit before proceeding to the next. The most comprehensive review of PSI effectiveness remains a meta-analysis documented by Kulik, C-L Kulik, and Cohen (1979a) who reported that the 75 PSI programs they studied raised final examination scores by about 0.50 standard deviations over programs using conventional (non-PSI) means of instruction. They also found that PSI produced less variation in achievement, higher student ratings, fewer course withdrawals, and that these favorable results occurred across a variety of subject matters and course settings. However, Keller, writing in 1985, was pessimistic about the future of PSI. He cited the large investment of instructor time needed to set up PSI courses and the general lack of support from university administrators as especially problematic. His concerns are corroborated by Lloyd and Lloyd (1986) who reported that progressively fewer PSI courses are being taught and many of those depart substantially from the recommended PSI format.

Table 1. Effect Sizes of Non-ICW Systems for Individualizing Instruction

System	Effect Size	N
Personalized System of Instruction (Keller's PSI)	.50	75
Audio-Tutorial (Postlethwait's A-T)	.20	42
Programmed Instruction		
Elementary-Secondary Mathematics	.11	89
Secondary Education (Various Subject Matters)	.08	47
Higher Education (Various Subject Matters)	.24	57

Audio-tutorial approach. The audio-tutorial approach (A-T) resembles PSI in that it also modularizes instructional content into units, is mostly applied in higher education, and leaves much of the individualization up to the students. Its basic form was developed in the early 1960s by Samuel Postlethwait and later described by Postlethwait, Novak, and Murray (1972). It consists of individual study sessions using audiotapes and/or other self study media, weekly group assembly sessions for lectures, films, and major examinations, and small group (6-10 students) quiz sections. As with PSI, the most thorough assessment of A-T was performed by Kulik and his associates (Kulik, C-L Kulik, and Cohen, 1979b). In summarizing the results of 42 studies, Kulik et al. reported that A-T increased overall student achievement by about 0.20 standard deviations over conventional (non-A-T) means of instruction. This finding held for a variety of subject matters and higher education settings. Thus the overall positive impact of A-T appears to be genuine, but small.

Programmed instruction. Although the development of programmed instruction was directly influenced by B.F. Skinner and his seminal paper on "The Science of Learning and the Art of Teaching" (Skinner, 1954), most implementations of this approach are closer to the intrinsic programming approach described by Crowder (1962) and commonly seen in programmed textbooks. Hartley reviewed results from 89 studies of elementary and secondary school mathematics instruction and reported an average improvement of 0.11 standard deviations through the use of programmed instruction. In a review of 47 comparisons, C-L Kulik, Shwalb, and Kulik (1982) reported that programmed instruction used in secondary education to present a variety of subject areas improved performance by about 0.08 standard deviations. In a review of 57 studies, Kulik, Cohen, and Ebeling (1980) reported that programmed instruction used in higher education to present a variety of subject areas improved performance by about 0.24 standard deviations over conventional instruction. Overall these results suggest that the impact of programmed instruction, like that of A-T, is positive, but limited.

Effect sizes for computer-based instruction are shown in Table 2. Again, these results are primarily based on meta-analyses performed by Kulik and his colleagues. Earlier evidence from 'box-score' studies on the effectiveness of computer-based instruction is also available. One study was performed by Vinsonhaler and Bass (1972), who found a median student increase in achievement of about 40% for computer-based instruction (CBI) compared with more conventional approaches. Another study by Orlansky and String (1979) found that the effectiveness of computer-based instruction used for military training is about the same as conventional instruction.

As shown in Table 2, Kulik, C-L Kulik, and Bangert-Drowns (1985) found an increase of 0.47 standard deviations across 28 studies of CBI used in elementary schools. Bangert-

Drowns, C-L Kulik, and Kulik (1985) found an increase of 0.40 standard deviations across 42 studies of CBI used in secondary schools. C-L Kulik and Kulik (1986) found an increase of 0.26 standard deviations across 101 studies of CBI used in colleges. At this point Kulik and his colleagues were prepared to conclude that the older the student, the less effective CBI was likely to be over conventional instruction. However, Kulik, C-L Kulik, and Shwalb (1986) found an average increase of 0.42 standard deviations across 24 studies of CBI used in adult education.

Kulik and his colleagues have not been the only analysts of CBI effectiveness. There have been enough reviews of CBI effectiveness that the reviews themselves have been summarized by Niemiec and Walberg (1987) who concluded that CBI raises achievement by about 0.42 standard deviations over conventional instruction.

Table 2. Effect Sizes for Computer-Based Instruction

Setting	Effect Size	N
Elementary School	.47	28
Secondary School	.40	42
Higher Education	.26	101
Adult Education	.42	24

Fletcher (1990) reported the findings shown in Table 3 for interactive videodisc (IVD) instruction. These suggest an improvement of 0.51 standard deviations across 9 studies of IVD used in industrial training, 0.69 standard deviations across 14 studies in higher education settings, and 0.39 standard deviations across 24 studies in military training. Although .39 is smaller than the effects of IVD instruction seen in the other two settings it is still equivalent (roughly) to raising the performance of 50th percentile students to the 65th percentile of performance. Its smaller size may be due to the larger standard deviations of performance and ability observed in populations of military students and/or the practice in military instruction of releasing students from school once they have crossed threshold levels of performance. The average effect size observed across all 47 of these studies is 0.51, which is roughly equivalent to raising the performance of 50th percentile students to 69th percentile performance.

Table 3. Effect Sizes for Interactive Videodisc Instruction

Setting	Effect Size	N
Military Training	.39	24
Industrial Training	.51	9
Higher Education	.69	14

In the case of ICW programs, it is particularly interesting to examine the impact of interactivity. This issue is addressed in Table 4 which shows effect sizes from single evaluation experiments intended to compare different levels of activity within the same IVD materials. All six studies suggest that interactivity contributes to student achievement in interactive videodisc instruction.

Table 4. Value of Interactivity in Interactive Videodisc (IVD) Instruction

Comparison	Effect Size Difference Favoring Interactivity	
More IVD Instruction Versus Less IVD Instruction	.14	(Experiment 1)
	.12	(Experiment 2)
Branched IVD Instruction Versus Linear IVD Instruction	.85	(Experiment 1)
	.59	(Experiment 2)
Interactive IVD Instruction Versus Linear IVD Instruction	1.54	
Level III IVD Instruction Versus Level II IVD Instruction	.46	

The range of positive results for ICW programs, suggest that in many settings and subject matters they are more effective than conventional approaches to instruction and that their use should be pursued. However, the essence of administrative decision making is deciding what must be given up to achieve new, beneficial results. The story is incomplete without consideration of ICW program costs.

IV. The Costs of Interactive Courseware

A. Issues in Measuring the Costs of Interactive Courseware

Levin (1983) has provided a primer on the measurement of costs in instruction and Knapp and Orlansky (1983) have developed a detailed cost model for training. These discussions rely on and refer to numerous earlier efforts that represent fairly well the state of the art.

Costs reported for any approach will vary widely depending on the assumptions and procedures used to collect -- or ignore -- different costs. If these assumptions and procedures are not explicit, it becomes difficult or impossible to judge the relevance of the cost data and to comprehend the conclusions that are based on them. Budgets are inadequate for cost estimation because they do not include all the components needed; they may distort the true costs of a component to comply with local accounting practices; and they represent planned, not actual expenses. After-the-fact analyses of costs are inadequate for roughly the same reasons, and it is especially difficult to disaggregate costs for one purpose (e.g., costing) after they have been assembled for another (e.g., production control or auditing).

One requirement in measuring the costs of instruction is a practicable list, or structure, of well defined cost components. These are provided in both the Levin and Knapp and Orlansky models. They are called ingredients by Levin and elements by Knapp and Orlansky. These lists should and apparently do capture all the components that are needed for cost-effectiveness comparisons of instructional alternatives regardless of the scope, complexity, or technology of the alternative. They also ensure a level of detail in the analysis that clearly identifies cost drivers -- the major contributors to the cost of the alternative. They can be used to select, plan, assess, and modify instructional alternatives.

Conceptually, the cost of an alternative consists of those resources that must be sacrificed in order to implement it. Since these resources could be used for other things, we speak of costs in terms of opportunities given up to secure the opportunities we have chosen. This notion of opportunity costs is the basis for cost-effectiveness analysis.

Levin suggested five classes of ingredients to be considered in a cost model. These are: Personnel, Facilities, Equipment and Materials, Other Program Inputs, and Client Inputs. Personnel costs include all the resources required for the human resources needed by the approach. Levin recommends that all personnel be classified according to their roles (instructional, administration, clerical, etc.), qualifications (training, experience, specialized skill), and time commitments (full time, part time). Facilities costs include all resources required to provide physical space for the approach. Since we are concerned with opportunity costs, these resources should be included whether paid for by the approach's implementors or not. Equipment and materials include furnishings, instructional equipment, and supplies -- again, whether paid for by the implementors or not. Other inputs include components that do not fit elsewhere. Examples of these are instructor training sessions and insurance costs. Client inputs include resources that must be contributed by the users (most commonly students and/or their employers) of the instructional approach. Client inputs are especially relevant in military and industrial training where student pay and allowances may be provided by the client.

Four categories generally found in cost models are: Research and Development, Initial Investment, Operating and Support, Disposal and Salvage. Research and development costs consist of all hardware, software, other materials, people, and facilities necessary to create, test, and evaluate an instructional approach. Initial investment costs comprise the one-time costs of procuring and deploying resources in the quantities needed to satisfy anticipated requirements for an instructional approach. Operating and support costs include those needed for managing, operating, and maintaining an instructional approach after it has been implemented. Disposal and salvage costs comprise the one-time costs of removing the instructional approach from operational use.

The Knapp and Orlansky model uses the first three of these categories to provide a common basis for cost evaluations and cost comparisons of all military training programs. The Knapp and Orlansky model is composed of the following elements:

A. <u>Research and Development</u>

1. Design
2. Component Development
3. Producibility Engineering and Planning
4. Tooling
5. Prototype Manufacturing
6. Data
    a. Managerial
    b. Technical

7. Test and Evaluation
8. System/Project Management
9. Facilities
10. Other

B. Initial Investment

1. Production
    a. Nonrecurring
        i. Production Planning
        ii. Production Tooling and Equipment
        iii. Industrial Facilities
        iv. Others
    b. Recurring
        i. Manufacturing
        ii. Sustaining Engineering
        iii. Sustaining Tooling
        iv. Quality Assurance
        v. Other
    c. Initial Spares and Repair Parts
2. Engineering Changes
3. Purchased System-Peculiar Equipment
4. Common Equipment
5. Data
    a. Managerial
    b. Technical
    c. Instruction Materials
6. Test and Evaluation
7. System/Project Management
8. Rents
9. Operational/Site Activation
10. Initial Training
    a. Instructors
    b. Maintenance Personnel
11. Transportation
    a. First Destination
    b. Second Destination

12. Other

C. <u>Operating and Support</u>

1. Direct Costs
    a. Instructional Costs
        i. Pay and Allowances
            (a) Instructors
            (b) Supervisors, Administration, and Support Personnel
            (c) Maintenance Personnel
        ii. Other Government Personnel Costs
        iii. Consumption
            (a) POL (Petroleum, Oil, and Lubricants)
            (b) Training Munitions
            (c) Utilities
                (i) Electric Power
                (ii) Other
            (d) Instructional Materials
            (e) Other
        iv. Replenishments and Spares
        v. Modification Materiel
        vi. Depot Maintenance
            (a) Labor and Materials
            (b) Second Destination Transportation
            (c) Other
        vii. Other Purchased Services
        viii. Other
    b. Training Activity Costs
        i. Pay and Allowances
        ii. Other Governmental Personnel Costs
        iii. Other
    c. Airfield and Carrier Operations Costs
        i. Pay and Allowances
        ii. Other Government Personnel Costs
        iii. Other
    d. Student Costs
        i. Pay and Allowances

ii. Other Student Costs

e. Other Direct Costs

2. Indirect Costs

a. Base Operations

i. Pay and Allowances

ii. Other Government Personnel Costs

iii. Other

b. Inventory and Supply Management

i. Pay and Allowances

ii. Other Government Personnel Costs

iv. Other

c. Military Family Housing Support

i. Pay and Allowances

ii. Other Government Personnel Costs

iii. Other

d. Command Support Costs

i. Pay and Allowances

ii. Other Government Personnel Costs

iii. Other

e. Other Indirect Costs

This model involves 11 elements for research and development, 25 elements for initial investment, and 39 elements for operating and support -- 75 elements in all. However, some of these elements can always be eliminated in estimating the costs of specific alternatives because they will be irrelevant -- munitions costs might be ignored in training for cooks. Other elements can be eliminated from specific cost-effectiveness evaluations because the costs will be the same across all alternatives -- command support costs may be the same in a comparison of tank gunnery trainer devices. The Knapp and Orlansky model appears to be both comprehensive and practicable for most military training applications and is a candidate for wide adoption.

B. What Do We Know about the Costs of Interactive Courseware?

Research and development costs are difficult to collect, rarely reported, and their place in cost-effectiveness analyses is unclear. Salvage and disposal costs are not as difficult to collect and their use in cost-effectiveness analyses is clearer, although likely to be slight, but no instance has been found where they have been reported or considered in analyses of instructional alternatives. Neither of these cost categories are considered further here.

This exclusion leaves initial investment costs and operating and support costs as categories under consideration. In some studies these two categories are combined; others allow them to be separated. Although both categories suggest favorable arguments for ICW programs, the case for ICW is usually stronger when operating and support costs rather than initial investment costs are considered.

The available cost analyses are usually based on cost ratios in which the costs of the ICW program are divided by the costs of the comparison, conventional instruction. The smaller the resulting ratio, the more favorable the cost argument for the ICW program.

Orlansky and String (1979) prepared the first and now classic general review of CBI costs and effectiveness. Although they found cost data to be limited and incomplete, they concluded that the cost of CBI was less than conventional instruction due to an average 30 percent savings in student time. No general review of CBI costs was found to have been completed since this one. More cost data on CBI have become available in the interim, and a new review of CBI costs may now be a realistic and worthwhile goal.

Fletcher (1990) reported cost ratios for IVD instruction evaluations that considered both costs and effectiveness. All 13 cost ratios found in studies of this sort were less than 1.00, indicating lower costs in each measured instance for IVD instruction. Five of these cost ratios were for initial investment and averaged 0.43. The remaining 8 cost ratios were for operating and support costs and averaged 0.16.

Walker (1985) documented an industrial training study in which the costs of delivering interactive videodisc instruction to remote sites were compared with the costs of centralized training. When the initial investment costs for developing and installing the interactive videodisc training were amortized over 3 years, Walker found that the costs per student were $1,568 for the centralized training and $553 for the interactive videodisc training -- a cost ratio of 0.35 for combined initial investment and operating and support costs.

Maher (1988) completed an extensive study on the feasibility and costs of five methods for training smog-check mechanics for the State of California. He concluded that because of the elimination of instructors and reduced training time, videodisc instruction would provide the most cost-effective approach of five that were considered for hands-on mechanic training and verification testing. Maher found that the costs for videodisc training would be $50.60 per student compared with baseline costs of $102.78 per student -- a cost ratio of 0.49 for combined initial investment and operating and support costs.

In summary, it appears that favorable cost ratios exist for IVD programs compared to conventional programs of instruction and that based on available but partial evidence, similarly favorable cost ratios may exist for CBI programs compared to conventional programs. Overall, then, favorable cost arguments can be made for the use of ICW programs, but available data do not support definitive or comprehensive conclusions.

V. The Cost-Effectiveness of Interactive Courseware

A. Issues in Measuring the Cost-Effectiveness of Interactive Courseware

The main issue in cost-effectiveness analyses of instructional programs appears to be that both cost and effectiveness data must be collected under the same empirical design using systematic models of cost inputs and effectiveness outputs. There are at least two reasons for this. First and most obviously, the same experimental controls should hold for all conditions (i.e., the experimental and control treatments) under which cost and effectiveness data will be drawn. Second, the same cost models should be used for all conditions and related to effectiveness measures in the same way across all conditions.

Kazanowski (1968) developed a standardized, 10-step approach to cost-effectiveness evaluations. Although the approach was oriented toward the selection of weapon systems, it provides a foundation for developing a general approach to evaluating cost-effectiveness in instruction. Kazanowski's 10 steps applied to instruction are the following:

(1) Define the objectives. In any systematic approach to instruction, an analysis must be performed to identify and define what the instruction is supposed to do -- to establish the instructional objectives. These objectives are most often expressed in terms of what students can do or attributes they possess once they finish the instruction. The objectives may be derived directly from the skills and knowledge required to perform a job, as they are in training, or they may be derived from national needs for a capable workforce and informed electorate, as they are in education. In the absence of these objectives, systematic

design, development, implementation, and evaluation of the instruction is impossible. It is, of course, possible to proceed non-systematically, and this is often done.

(2) Identify the mission requirements. As discussed above, it is fundamental that an instructional alternative meet its objectives at a micro level by bringing about the instructional outcomes we expect to see in its graduates. On a more macro level, the instructional alternative as a system must possess productivity, or 'pipeline', characteristics defined by its mission -- it must be able to turn out a given number of graduates within a given amount of time. These pipeline requirements are determined by the instructional mission.

(3) Develop the alternatives. Once the instructional objectives in terms of student outcomes are defined and once the pipeline requirements of the instructional program are established, alternative approaches for satisfying these objectives and requirements must be identified. As in most analyses intended to support decisions, the generation of alternatives is a critical activity requiring considerable imagination and creativity. There are tools and aids but no effective procedures for developing comprehensive sets of instructional alternatives to subject to cost-effectiveness evaluation.

(4) Design the effectiveness evaluation. The measures to be examined by an effectiveness evaluation should be well defined, evident, and established at an early stage. These measures should follow directly from the instructional objectives and pipeline requirements established for the program in the first and second steps above. The omission of significant measures can invalidate the results of an evaluation, but the inclusion of too many measures (perhaps more than 10) could paralyze final choice of an alternative. A test of the adequacy of the measures selected is to determine whether an instructional alternative could excel in most of the measures and still not be best on some intuitive level. If this is true, then some important measures are missing.

(5) Select a fixed cost or fixed effectiveness approach. A key distinction between education and training is that training is a means to an end, education may be viewed as a valuable end in its own right. In training we may hold effectiveness constant and try to minimize cost to achieve a required level of effectiveness, whereas in education we may hold cost constant and try to maximize effectiveness to be obtained for that cost. The choice between fixed cost and fixed effectiveness approaches may depend on whether the client decision maker is a trainer or an educator.

(6) Determine the capabilities of the alternative systems. Once the measures and evaluation approach have been settled, it is time to gather the data and proceed with the analysis. This should be done through the development and implementation of a systematic empirical design that will provide reliable answers to the questions being addressed.

(7) Tabulate the alternatives and measures. When the data are gathered, they should be tabulated in a form suitable for comparison. In a fixed cost approach, the alternatives are usually tabulated following the 'Northwest' rule. That is to say, the measures are listed from left to right in decreasing priority, and the alternatives are listed from top to bottom in order of decreasing apparent value.

(8) Analyze the merits of the alternative systems. The client decision makers may feel that a tabulated array of findings resembles raw data more than useful information. Discussion of the findings, the strengths and weaknesses of the data collection procedures, and assessment of the alternatives in light of the evaluation are in order and should be provided.

(9) Perform a sensitivity analysis. The outcome of a cost-effectiveness evaluation may or may not be sensitive to the assumptions on which it is based. The sensitivity analysis is intended to find this out. Cost-effectiveness evaluations in instruction may be sensitive to assumptions concerning personnel costs, software maintenance costs, costs of consumables (e.g., fuel and ammunition in military training), mean times between failure for simulators and actual equipment, actual equipment operation costs, and student-instructor ratios.

(10) Document the bases of the previous nine steps. No cost-effectiveness evaluation will be perfect. It is critical for decision makers to know its strengths and limitations. It is not a trivial task to identify the assumptions und lying an evaluation, and these assumptions should be identified and described a icitly as possible. The underlying models of cost and effectiveness should be documented so that decision makers can see what has been excluded and therefore assumed either irrelevant or equivalent across alternatives.

B. What Do We Know about the Cost-Effectiveness of Interactive Courseware?

The Orlansky and String (1979) study discussed earlier remains the primary source of information on the cost-effectiveness of CBI in military training. It reported that (a) student performance under CBI was at least as good as under conventional instruction and (b) students reached the desired levels of performance in about 30 percent less time using CBI. Although these result suggest a favorable argument for the cost-effectiveness of CBI, Orlansky and String did not draw a conclusion on its cost-effectiveness due to the limited and incomplete nature of both cost and effectiveness data.

Spindler (1990) reported results for CBI used in maintenance training that were similar to those of Orlansky and String (1979). He reported that there were no statistical differences in measured performance of CBI and conventionally trained students on either end of training measures or on measures of retention in four different maintenance courses. Additionally, he reported an average 33 percent time savings for the CBI students to reach criterion performance which resulted in an average 19 percent cost savings for the CBI groups.

It may be that the 30 percent time savings will remain stable across many different forms of ICW programs. Fletcher (1990) also reported a student time savings of about 30 percent in his review of IVD instruction. Because IVD costs were generally lower than those reported for conventional instruction in the studies reviewed and because IVD effect sizes were generally higher than those reported for conventional instruction, he reported a "suggestive" but not conclusive finding that IVD instruction is more cost-effective than conventional instruction based on the studies included in this review.

A CBI study reported by Fletcher, Hawley, and Piele (1990), reports cost-effectiveness data from an empirical study of CBI mathematics instruction given to 3rd and 5th grade Canadian children and compares these results with earlier ones reported by Jamison, Fletcher, Suppes, and Atkinson (1975) and with after-the-fact analyses of cost-effectiveness performed by Levin, Meister, and Glass (1987). Effectiveness measures in all these cases are grade placements from standardized test scores on scales described as "Total Mathematics." Cost and effects from these studies are shown in Table 5. Costs per student per year are based on a daily average of 10 minutes of CBI for a 180 day school year. The costs of reducing class size were estimated by Levin et al. for mathematics instruction only -- they are not intended to represent the costs of reducing class size for all instruction. The cost-effectiveness ratios reported in all cases are the annual costs (in 1985 dollars) to raise the control group's achievement by one standard deviation.

Table 5. Cost-Effectiveness of CAI in Mathematics from Three Studies

Instructional Alternative	Cost[a] Per Student Per Year	Mean Effect Size (SD Units)	Cost /Effect Ratio
Jamison et al. (1976)			
CBI Grade 3 -- 10 Minutes Per Day			
CA Students	$250	.74	$338
MS Students	250	1.23	203
CBI Grade 5 -- 10 Minutes Per Day			
CA Students	250	.54	462
MS Students	250	.51	490
Levin et al. (1987)			
Tutoring			
Peers 20 Minutes Per Day	277	.97	286
Adults 20 Minutes Per Day	1080	.67	1612
CBI 10 Minutes Per Day	143	.12	1192
Reduce Class Size			
From 35 to 30	59	.06	983
30 to 25	82	.07	1171
25 to 20	123	.09	1367
35 to 20	263	.22	1195
Increase Instructional Time			
30 Minutes Per Day	80	.03	2667
Fletcher et al. (1990)			
CBI Grade 3 -- 10 Minutes Per Day	46	.24	192
CBI Grade 5 -- 10 Minutes Per Day	37	.18	206

[a]All costs are reported as 1985 US Dollars.

Fletcher et al. conclude from their findings that computer-based instruction substituted for some portion of traditional instruction may be a superior cost-effective approach for raising total mathematics scores as measured by standard tests when compared with tutoring, reductions in class size, increased instruction time, and earlier CAI that did not employ stand-alone microcomputers. The closest competitor appears to be peer tutoring. The best practice may be to combine the best features of both approaches and present CAI to two (or more) students at a time at a single student station.

Pairing students should provide many of the benefits of one-on-one peer tutoring, most of the benefits of CAI, and double the cost-effectiveness of CAI. This approach echoes back to some of the earliest experimental investigations of CAI. In 1964, Grubb reported that low-ability students studying statistics who were paired at computer terminals achieved posttest scores that were as good as those of high-ability students who worked at the terminals alone. Okey and Majer (1976) found no significant differences in the achievement of college students who worked singly, in pairs, and in groups of 3-4 for three hours at CAI workstations. Dossett and Hulvershorn (1983) found no significant differences in the achievement of Air Force personnel who worked singly and in pairs during a week-long block of instruction at CAI workstations. Mevarech, Stern, and Levita (1987) found no significant differences in the achievement of junior high school students working individually or in pairs at CAI workstations. Shlechter (1988) found no significant differences in the achievement of soldiers working individually, in groups of 2, or in groups of 4 at CAI workstations. If achievement does not suffer, then the lower cost of grouping students on CAI workstations argues strongly for the cost-effectiveness of this approach.

In brief, there is evidence that in many settings and for many subject matters ICW programs are more effective and less costly than more conventional approaches to instruction, but conclusive evidence on the cost-effectiveness of ICW programs remains to be obtained.

Whatever the empirical results that emerge, cost-effectiveness will still have to be decided on a case by case basis. It may not be fair to ask about the cost-effectiveness of ICW programs without knowing what alternatives are being compared in what settings, for what subject matters, and to achieve what instructional objectives.

On the other hand, more general results may be achieved by assessing the cost-effectiveness of a variety of instructional alternatives in achieving specific classes of instructional objectives derived from a specific model of instruction. For instance, based on the Instructional Quality Profile proposed by Merrill, Reigeluth, and Faust (1979), we might assess and provide general findings on the cost-effectiveness of different instructional alternatives used to prepare students to recall a fact, apply a concept, recognize a procedure, or recall a principle.

VI. Utility Analysis

The problem of choosing among alternatives may involve more than the determination of relative effectiveness and costs. All instructional outcomes may not be equal in the view of decision makers, and choices among instructional alternatives may be guided by the weights they assign to different instructional outcomes. Utility analysis takes these weights into account by assigning a numerical value based on the desirability of each instructional outcome. When utility measures are combined with cost-effectiveness analysis, the resulting decision process more closely reflects decision makers' values.

In this approach, the value or utility of each outcome is measured by asking several respondents (decision makers, or their representatives) to rate it on some scale and then by aggregating the ratings to obtain an overall utility. The respondents chosen to make the ratings should be those who have a stake in the choice of alternatives. Field commanders, training policy makers, or training managers may be best for military training. The primary methodological issues are to ensure that the definitions of each outcome are clearly articulated to all respondents and that the respondents chosen are appropriate for answering the question being asked.

Once chosen, the respondents are asked to assume that some numerical value, say 10, had been assigned to one of the outcomes. They are then asked to assign -- on a scale of 0 to perhaps twice the baseline value chosen, say 20 -- proportionally higher or lower ratings to each of the other outcomes. As examples and guidelines, they are told that a rating half as large as the baseline, say 5, means that the outcome is one-half as valuable to them as the first outcome chosen and that a rating half larger than the baseline, say 15, means that the outcome was one and one-half times as valuable to them as the first outcome chosen. The average utility value assigned to each outcome is then calculated. This interview-based approach follows procedures advocated by Stokey and Zeckhauser (1978).

The cost utility ratio of each treatment is calculated by applying the following formula:

$$\text{Cost-utility} = \frac{C}{\Sigma\,(U_i X_i)}$$

where

C is the cost per student of the instructional approach,
X_i is the mean change in the measure associated with outcome i,
U_i is the utility of X_i as assessed by interviewees.

The lower the cost utility ratio, the greater the cost-utility of the alternative.

All utility analysis does is apply weights to the various outcomes in a cost-effectiveness analysis. In the above example, the weights were assigned by decision makers. They could of course be assigned by an amount that is determined empirically -- by the contribution each outcome makes to effectiveness in performing some military mission. At present these empirically derived measures are not available, and interviewing responsible decision makers may be the best current approach for incorporating utility analysis in our cost-effectiveness analyses.

VII. Issues in Assessing the Cost-Effectiveness of Interactive Courseware

A. Effectiveness

Among unsettled issues in the assessment of ICW program effectiveness are the following:

- Third party evaluation. Many evaluations in military training settings were performed by the developers who also produced the ICW programs being evaluated -- few evaluations were performed by third parties. There are strengths and weaknesses in evaluations performed by developers. Developers are rarely indifferent to the success of their products and may, intentionally or not, bias the results of their evaluation. On the other hand, developers have a stake in honest assessment, and they may understand better than anyone -- including potential users -- the strengths and limitations of what they have produced. In any case, we should seek to understand better the implications of developer versus third party evaluations since this practice is so common in the evaluation of military ICW programs.

- Uniqueness. One difficulty for any evaluation of an innovative technology is that there is nothing else like it. Each new technology has its own strengths and limitations. If the evaluation is held to strict instructional and experimental controls based on the older technology, the newer technology will be at a disadvantage. Also, new approaches are unlikely to be used well since not enough is understood about how best to employ them. The early horseless carriages was certainly inferior in both costs and effectiveness to horse drawn carriages if viewed strictly as a means for getting from one place to another. Similarly, we may have incipient in the functionalities of ICW programs, promise and applications of which we are only vaguely aware. Despite all the evaluations listed, we may have yet to see an instructional approach that uses ICW technology to best advantage.

- Redesign of materials. Another perennial problem in comparison studies of new approaches with conventional approaches is that the content and objectives of the conventional approach may be revised and incorporated in the new but not the conventional approach. The revised body of materials may then be compared with the original conventional instruction, owe more of its success to the revision than to the functionalities of the new approach. More effort may be needed to avoid or work around these paradigms in the evaluation of ICW programs.

- Production quality. Direct assessment of ICW content or production quality was beyond the scope and resources of most of the studies reported here. Aspects such as the quality of graphics, clarity of instructional text, verisimilitude of simulations, and relevance of tutorial advice may have had a substantial impact on the effectiveness of many ICW programs, but these issues were generally not addressed. The impact of production quality on both the costs and effectiveness of ICW programs should be better understood.

- Relation of instructional design to outcomes. Different outcomes, or instructional objectives, must compete for scarce instructional resources. Decisions made in the design of ICW programs impact both their costs and their achievement of specific instructional objectives. These relationships should be better understood. How, for instance, should we design a program to maximize transfer ability, speed of response, or motivation to continue study? What do these designs cost? To what degree do they contribute to instructional effectiveness? How do we trade them off against one another, as we invariably must in the practical world of training? The individualization of control and student progress that we can exercise in ICW programs, raises these issues to a level of both significance and practical payoff that they do not reach elsewhere.

- Linkage to job performance. We need to know better how instructional outcomes are related to the skills and knowledge required to perform military jobs. Instruction design is often assumed to fall easily, perhaps trivially, out of comprehensive analyses of the skills and knowledge needed to perform military jobs. This rarely occurs in practice. Our measures of these skills and knowledge, our ability to assess them in individuals performing the job, our ability to link them to instructional outcomes, our understanding of how to weigh them in designing a program of instruction are all limited and incomplete. Again, these issues were much less important when we had less control over the individualization of control and student progress than we do now with the advent of ICW technology.

B. Cost

Among unsettled issues in the assessment of ICW program costs are the following:

- Adoption of a comprehensive cost model. Perhaps the most obvious and pressing issue in cost-effectiveness analysis of any instructional program is the need to adopt a comprehensive cost model that includes carefully defined definitions of all its elements so that cost analyses produce the same results in all hands, comparisons can be made across different studies, and data on instructional approaches can be shared by all concerned. The Knapp and Orlansky (1983) cost model is an excellent beginning, but if this model is not generally accepted and adopted some other one should be.

- Baseline measures. Discussions of innovative approaches such as ICW tend to focus on the new and neglect the current. To understand the cost impact of ICW programs we need to better understand the costs of what we do now. The costs of conventional approaches are frequently neglected in studies of ICW cost-effectiveness or they are simply not available. We should place more emphasis on capturing the costs of conventional instruction approaches while they are readily available and including them in accessible databases.

- Database of costs. The costs of both conventional and innovative approaches to instruction, collected in accordance with a commonly accepted model, should be maintained in an accessible, easily usable database made up of commonly defined, well understood cost data elements.

C. Cost-Effectiveness

- Media selection database. When should we use ICW programs? This question immediately expands to what combination of media and then what combination of instructional approaches should we use to achieve our training objectives? Finding this optimal combination -- or even a good combination -- of instructional approaches depends on the availability of a comprehensive set of data elements describing their costs and effectiveness. At the very least, then, a sharable database on costs and effectiveness with a well-defined, comprehensive set of data elements should be prepared and made available to all users. Development of such a database could begin almost immediately, and it would be of immediate value to all military organizations. Given this database, media selection might proceed simply through the use of an automated tool such as AIMS (Automated Instructional Media Selection) (Kribs, Simpson, and Mark, 1983) which in its turn would achieve sufficient practical utility to be widely used and supported by training developers.

- Need for data. Collection of cost and effectiveness data in the evaluation phase of any development is too infrequent. It should be done more frequently, even routinely. As suggested above, evaluations that address effectiveness only or costs only have their place, but when instructional decision makers make choices in the practical world of training, they need data on both. These data should be more widely available.

- Full system application. The development of ICW programs involves all the standard phases of system development -- analysis, design, production, implementation, and evaluation. Most cost-effectiveness assessments of ICW programs focus on the impact of the final product, few of these assessments attend to the cost-effectiveness of ICW development. The use of ICW is decided in the design phase, so the most significant cost-effectiveness analyses from the standpoint of its development may be for the production, implementation, and evaluation of ICW programs. ICW production has received some attention in that the productivity of different authoring systems or languages continues to be assessed (e.g., Seidel and Park, in press), but more assessment is needed. For instance in-house versus contract production of ICW programs is a perennial issue in military training for which no substantive cost-effectiveness assessment could be identified. Additionally, no cost-effectiveness assessments of implementation alternatives and evaluation alternatives for ICW programs could be identified.

- Cost-effectiveness measures. A metric should be developed for combining effect size with cost so that if costs become more frequently available from evaluations of ICW programs, we will be able to quantitatively aggregate results from cost-effectiveness assessments in much the same way that meta-analysis allows us to quantitatively aggregate results from many, separate effectiveness assessments.

- Cost-utility. Cost-utility assessment is conceptually and procedurally simple, and it is needed by instructional decision makers. Serious consideration should be given to its inclusion in cost-effectiveness assessments.

VIII. General Recommendations

General recommendations for assessing the cost-effectiveness of training have changed little over the last 5-10 years. Orlansky (1992) provided a succinct review of these recommendations. His four recommendations were:

- Include trade-offs between costs and effectiveness in cost-effectiveness evaluations. Especially show how effectiveness varies with cost and how improved methods of instruction can increase effectiveness without increasing costs.

- Include factors of learning and forgetting in cost-effectiveness evaluations. Effectiveness in general is not solely assessed by skills and knowledge achieved at a single point in time, but also by how rapidly these are achieved and how rapidly they decay over many points in time.

- Develop data bases on costs, effectiveness, and cost-effectiveness of training systems so that the costs and effectiveness of new training systems can be estimated before they are built.

- Give high priority to investigations of military effectiveness. There are many paths to military effectiveness. They include training, personnel selection, materiel acquisition, equipment design, supply schedules, among others. All these contributors to effectiveness carry a cost and all contribute in some way to effectiveness. A comprehensive technology of military operations would provide trade offs that maximize effectiveness and minimize costs for a comprehensive set of missions.

  Development of such a technology would substantially improve military capabilities, but it is a long way from being accomplished. However, limited, pairwise trade offs between training and other contributors to military effectiveness may be within our grasp. We might trade off training against weapons capabilities, training against selection, training against equipment supply, training against materiel, and so forth to determine optimal combinations that would maximize effectiveness and minimize costs for the variables under consideration. These trade offs could be applied during acquisition cycles and evolve into a concurrent engineering of human performance.

These issues remain valid and deserving of attention. Military effectiveness has benefited from many analytical techniques. Cost-effectiveness analyses will allow us to devise optimal balances in applying our increasingly scarce resources to meet the exacting human performance requirements of modern military operations. These analyses deserve high priority, continued development, and wide application.

IX. References

Bangert-Drowns, R.L., Kulik, J.A., Kulik, C-L.C. (1985) Effectiveness of computer-based education in secondary schools. Journal of Computer Based Instruction, 12, 59-83.

Clark, R. E. (1983). Reconsidering research on learning from media. Review of Educational Research, 53, 445-459.

Crowder, N.A. (1962). Intrinsic and extrinsic programming. In J.E. Coulson (Ed.), Programmed learning and computer-based instruction (pp. 58-66). New York, NY: John Wiley.

Dossett, D.L., & Hulvershorn, P. (1983). Increasing technical training efficiency: Peer training via computer-assisted instruction. Journal of Applied Psychology, 68, 552-558.

Fletcher, J.D. (1990) The Effectiveness of Interactive Videodisc Instruction in Defense Training and Education (IDA Paper P-2372). Alexandria, VA: Institute for Defense Analyses.

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

Glass, G.V. (1976) Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G.V., McGaw, B., & Smith, M.L. (1981) Meta-analysis in social research. Beverly Hills, CA: Sage Publications.

Grubb, R.E. (1965) The effects of paired student interaction in the computer tutoring of statistics. Paper presented at the annual meeting of the National Society for Programmed Instruction, Philadelphia, PA.

Hartley, S.S. (1977) Meta-analysis of the effects of individually paced instruction in mathematics (Unpublished doctoral dissertation). Boulder, CO: University of Colorado. University Microfilms No. 77-29,926)

Hedges, L.V. and Olkin, I. (1985) Statis cal Methods for Meta-Analysis. Orlando, FL: Academic Press.

Jamison, D.T., Fletcher, J.D., Suppes, P., & Atkinson, R.C. (1976) Cost and performance of computer-assisted instruction for education of disadvantaged children. In J.T. Froomkin, D.T. Jamison, & R. Radner (Eds.), Education as an industry (pp. 201-240). Cambridge, MA: Ballinger Publishing Company.

Kazanowski, A.D. (1968) A standardized approach to cost-effectiveness evaluations. In: J.M. English (Ed.) Cost-Effectiveness: The Economic Evaluation of Engineered Systems. John Wiley, New York, New York, pp. 113-150.

Keller, F.S. (1968) Goodbye, teacher .... Journal of Applied Behavior Analysis, 1, 79-89.

Keller, F.S. (1985) Lightning strikes twice. Teaching of Psychology, 12, 4-8.

Knapp, M. I., & Orlansky, J. (1983) A Cost Element Structure for Defense Training (IDA Paper P-1709). Institute for Defense Analyses, Alexandria, Virginia.

Kribs, H.D., Simpson, A.C., and Mark, L.J. (1983) Automated Instructional Media Selection (AIMS) (Technical Report No. 79-C-0104-1). Orlando, FL: Navy Training Eq... ent Center.

Kulik, C-L.C., Kulik, J.A., and Shwalb, B. (1986) Effectiveness of computer-based instruction in adult education. Journal of Educational Computing Research, 2, 235-252.

Kulik, C-L.C., Shwalb, B.J., & Kulik, J.A. (1982) Programmed instruction in secondary education: A meta-analysis of evaluation findings. Journal of Educational Research, 75, 133-138.

Kulik, J.A., & Kulik, C-L.C. (1986) Effectiveness of computer-based education in colleges. AEDS Journal, 19, 81-108.

Kulik, J.A., Kulik, C-L.C., & Bangert-Drowns, R.L. (1985) Effectiveness of computer-based education in elementary schools. Computers in Human Behavior, 1, 59-74.

Kulik, J.A., Cohen, P.A., & Ebeling, B.J. (1980) Effectiveness of programmed instruction in higher education: A meta-analysis of findings. Educational Evaluation and Policy Analysis, 2, 51-64.

Kulik, J.A., Kulik, C-L. C., & Cohen, P.A. (1979a) A meta-analysis of outcome studies of Keller's personalized system of instruction. American Psychologist, 38, 307-318.

Kulik, J.A., Kulik, C-L. C., & Cohen, P.A. (1979b) Research on audio-tutorial instruction: A meta-analysis of comparative studies. Research in Higher Education, 11, 321-341.

Levin, H.M. (1983) Cost effectiveness. Beverly Hills, CA: Sage Publications.

Levin, H. M., Glass, G. V., & Meister, G. R. (1987) Cost-effectiveness of computer-assisted instruction. Evaluation Review, 11, 50-71.

Lloyd, M.E., & Lloyd, K.E. (1986) Has lightning struck twice? Use of PSI in college classrooms. Teaching of Psychology, 13, 149-151.

Maher, T.G. (1988) Hands-on verification of mechanics training: A cost-effectiveness study of videodisc simulation. Sacramento, CA: Bureau of Automotive Repair, California State Department of Consumer Affairs.

Merrill, M.D., Reigeluth, C.M., Faust, G.W. (1979) The instructional quality profile: A curriculum evaluation and design tool. In: H.F. O'Neil (Ed.), Procedures for Instructional Systems Development. Academic Press, New York, New York, pp. 165-204.

Mevarech, Z.R., Stern, D., Levita, I. (1987) To cooperate or not to cooperate in CAI: That is the question. Journal of Educational Research, 80, 164-167.

Niemiec, R., & Walberg, H. J. (1987) Comparative effects of computer-assisted instruction: A synthesis of reviews. Journal of Educational Computing Research, 3, 19-37.

Okun, A. M. (1970) The Political Economy of Prosperity. W W Norton, New York, New York.

Okey, J.R., & Majer, K. (1976) Individual and small-group learning with computer-assisted instruction. AV Communication Review, 24, 79-86.

Orlansky, J. (1992) NATO Studies on the Military Value and Cost-Effectiveness of Training. In K. Ellis, J. Moraal, and A. von Baeyer (Eds.) Workshop on the Military Value and Cost-Effectiveness of Training (Technical Proceedings AC/243 (Panel 7) TP/3). Brussels, Belgium: NATO Headquarters. NATO UNCLASSIFIED.

Orlansky, J., & String, J. (1979) Cost-Effectiveness of Computer Based Instruction in Military Training (IDA Paper P-1375). Institute for Defense Analyses, Alexandria, Virginia.

Postlethwait, S.N., Novak, J., & Murray, H.T., Jr. (1972) The Audio-Tutorial Approach to Learning. Minneapolis, MN: Burgess Publishing.

Seidel, R.J., & Park, O.K. (in press) Evaluation of CBT Authoring Systems: Need for a Data Base. In P.R. Chatelier and R.J. Seidel (Eds.) Advanced Technologies Applied to Training Design. New York, NY: Plenum Publishing.

Shlechter, T.M. (1988) The effects of small group and individual computer-based instruction on retention and on training lower ability soldiers (Research Report 1497). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Skinner, B.F. (1954) The science of learning and the art of teaching. Harvard Educational Review, 24, 86-97.

Spindler, G.B. (1990) Some Case Studies and Lessons Learned in Cost Effectiveness Analysis. In J.P. Kincaid, R. Ahlers, D. Carr (Eds.), Cost-Effectiveness Analysis of Training Systems Workshop: Proceedings. Orlando, FL: Institute for Simulation and Training, University of Central Florida.

Stokey, E., & Zeckhauser, R. (1978) A primer for policy analysis. New York, NY: W. W. Norton and Company.

Vinsonhaler, J.F., & Bass, R.K. (1972) A summary of ten major studies on CAI drill and practice. Educational Technology, 12, 29-32.

Walker, R.A. (1985) Videodisc for decentralized training. In Proceedings of the Seventh Annual Conference on Interactive Videodisc in Education and Training. Warrenton, VA: Society for Applied Learning Technology.