AD-A283 021

(1)

ARMY RESEARCH LABORATORY

Applying Training System Estimation Models to Army Training
Volume I: Analysis of the Literature

Frederick A. Muckler
Dorothy L. Finley

ARL-TR-463 May 1994

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1994	Final

4. TITLE AND SUBTITLE

Applying Training System Estimation Models to Army Training: Volume I. Analysis of the Literature

5. FUNDING NUMBERS

PR: 1L162716AH70
PE: 6.27.16

6. AUTHOR(S)

Muckler, F. A.; Finley, D. L.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Human Research & Engineering Directorate
Aberdeen Proving Ground, MD 21005-5425

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARL-TR-463

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A review and analysis of the literature about training system models and methods during the last 20 years was conducted. Thirty-six training system estimation models were identified. The models were evaluated with respect to the adequacy of their documentation and their ability to predict cost effectiveness, training efficiency, training effectiveness, and appropriate media selection. The purpose was to review the history and state of the art of such models and to identify their strengths and deficiencies. The goal was to see to what extent these models could be useful for enhancing future Army training system resource estimation early in the development of a materiel system.

14. SUBJECT TERMS

training system estimation models
training development methods
training resource estimation
training effectiveness prediction and evaluation
MANPRINT

15. NUMBER OF PAGES

40

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

One long-range research goal of the U.S. Army Research Institute, and subsequently the Human Research and Engineering Directorate of the U.S. Army Research Laboratory, is to develop methods for estimating the resources required to develop and implement the training system required to support a new materiel system. The objective is to influence the design of that materiel system early in its development program by showing the impact of alternate prime system designs on training costs. As an initial step in the development of such methods, a review and analysis of the literature about training estimation models and methods during the last 20 years was conducted. The purpose was to review the history and state of the art of such models and to identify their current strengths and deficiencies. The goal was to see to what extent these models could be useful for enhancing future Army training system resource estimation.

Thirty-six training system estimation models were identified. The definition of "system" was often limited in scope. With few exceptions, there is little connection or coordination among them except that they mostly provide some form of training and cost-effectiveness estimates. To a great extent, the last 20 years have been a period of trying ideas, some of which have been very complex. In one sense, the "state of the art" is a large learning experience from which many major future advances may be possible. One point of view is that the past two decades were necessary to structure and to begin to understand the problem.

This research has been performed under the auspices of Task 1209, Soldier-Equipment Considerations in Military Occupational Specialty (MOS) Design. Work under that area is focusing on the development of methods to optimize the task clusters, that is, task structure, assigned to MOSs when changes occur in doctrine or as a function of force modernization. One of the major impacts of changes in MOS task structure that must be accounted for is on the design of the training system supporting the MOSs. The research of MOS restructuring and this analysis of the training system estimation literature will provide inputs to the long-range goal of developing training system resource estimation methods that will affect materiel system design choices.

AMCMS Code 612716.H700011 Technical Report ARL-TR-463

APPLYING TRAINING SYSTEM ESTIMATION MODELS TO ARMY TRAINING: VOLUME I. ANALYSIS OF THE LITERATURE

Frederick A. Muckler
Dorothy L. Finley

May 1994

The Systems Research Laboratory of the U.S. Army Research Institute (ARI), referenced in this report, has been reorganized as part of the Human Research and Engineering Directorate of the U.S. Army Research Laboratory, effective 1 October 1992. The program and facility described here were completed before the organizational change.

APPROVED: Robin L. Keesee
ROBIN L. KEESEE
Executive, Human Research &
Engineering Directorate

Approved for public release; distribution is unlimited.

U.S. ARMY RESEARCH LABORATORY

Aberdeen Proving Ground, Maryland

CONTENTS

EXECUTIVE SUMMARY

RESEARCH REQUIREMENTS

The design of a prime materiel system largely determines the manpower, personnel, and training (MPT) resource requirements needed to support that system. One of the tenets of the manpower and personnel integration (MANPRINT) program states that to estimate the MPT resource requirements early in the prime system's development program increases the likelihood that prime system designs that are costly to support and noneffective will be avoided. That is, alternate system designs will be considered that are less MPT resource intensive.

One long-range research goal of the U.S. Army Research Institute, and subsequently the Human Research and Engineering Directorate of the U.S. Army Research Laboratory (ARL), is to develop methods for estimating the resources required to develop and operate the training system required to support a new materiel system. The objective is to influence the design of that materiel system early in its development program by showing the impact of alternate prime system designs on training costs.

PROCEDURE

As an initial step in the development of such methods, a review and analysis of the literature about training estimation models and methods during the last 20 years was conducted. The specific models that were identified were evaluated with respect to five criteria:

1. Does the model predict cost effectiveness?

2. Does the model predict training efficiency?

3. Does the model predict training effectiveness?

4. Does the model predict appropriate media selection?

5. Is there sufficient documentation to reconstruct the model?

The purpose was to review the history and state of the art of such models and to identify their current strengths and deficiencies. The goal was to see to what extent these models could be useful for enhancing future Army training system resource estimation.

FINDINGS

Thirty-six training system estimation models were identified. The definition of "system" was generally limited in scope. Many models, for example, concerned only training devices or other media and did not consider the overall training system (instructors, administrators, training location, facilities, etc.). The extent to which they could handle alternatives such as institutional versus unit training, collective versus individual training, and sustainment training needs and capabilities was unclear. None had been developed for the specific purpose of influencing prime materiel system design and were not sufficient for that purpose even though they presented potentially useful approaches. With few exceptions, there is little connection or coordination among them except that they mostly provide some form of training and cost-effectiveness estimates. Of the 36 models, seven afforded some, though different, capabilities with respect to all five criteria mentioned previously.

To a great extent, the last 20 years have been a period of trying ideas, some of which have been very complex. In one sense, the "state of the art" is a large learning experience from which many major future advances may be possible. One point of view is that the past two decades were necessary to structure and to begin to understand the problem.

USE OF FINDINGS

The findings from analyzing the literature will provide one input to the design of a research program to develop methods to estimate the development and implementation resource requirements for the training systems needed to support alternate prime materiel system designs. The current state of the art not only has deficiences, which represent potential research issues, but also many strengths, which should be exploited in this research program. Given the variety of concerns addressed and approaches taken by the reviewed 36 models, other researchers should deem our findings useful.

APPLYING TRAINING SYSTEM ESTIMATION MODELS TO ARMY TRAINING:
VOLUME I. ANALYSIS OF THE LITERATURE

INTRODUCTION

More than $200 billion per year are spent in the United States for institutional education and training. In the U.S. Army alone, the budget for institutional training exceeds $2 billion annually. It is of major concern to the training community that these sums be spent on effective training.

The major single development in education and training during the past generation has been the unparalled explosion of training technology. Based partly on the digital computer, new possibilities for training have escalated, allowing for training relationships with students never before achievable. At the same time, however, this new technology has been very expensive.

The new technology provides many options. The training system designer is not limited in knowledge instruction, for example, to the classic classroom, instructor-student, blackboard relationship. The designer may choose from many media, which have a very wide variety of learning situations. Available choices for knowledge instruction beyond the classic tools include such possibilities as intelligent tutoring systems, interactive video disks, and embedded training. But the question will be, what is (are) the best option(s) to select for any given specific training problem?

To aid the training system designer in making that choice, a set of tools called "training system estimation models" has been developed during the past 20 years. Thirty-six of these models have been identified through a literature search through such sources as the proceedings of the annual Human Factors Society conferences, Psychological Abstracts, the authors' personal libraries, and the National Technical Information Center (see Volume II). These 36 models are discussed here. Briefly, these models try to compare and evaluate, quantitatively, training system options. The scope of the training "system" definition covered by these models varies considerably. Some are limited to consideration of a training device, for example, while others consider development of, for example, all the institutional training needed to support a weapon system. They provide, for the most part, both training effectiveness and cost estimates. Their principal purpose is to aid in selecting the best training system option at the lowest reasonable cost before hardware and software commitments are made.

The purpose of this report is to review the history and the state of the art with respect to training system estimation models. Of interest are the current strengths and deficiencies of the models. The goal is to see to what extent these models could be useful for enhancing future Army training system development.

During the next decade, the Army will have many alternatives in the way training is used. First, there is no question that training technology will continue to expand and grow, particularly in all forms of computer-based training. Second, far more flexible and administratively complex training may be in the Army's future with reduced institutionalized training and increased training responsibilities and capabilities brought to units. The current concept of future Army training, termed "Army Training 21," calls for increased use of "distributed training," that is, initial and sustained training of knowledges and skills in the unit, as opposed to training in the institution. Third and finally, budgets for training will decrease rapidly as a function of a sharp drop in institutional training load driven by the significant downsizing of the Army. Since institutional training budgets are largely determined by the size of the student load (number of course hours multiplied by the number of students), there will be fewer resources for training and for training development. The move to distributed training and the increased criticality of reserve component and national guard unit capabilities will require more in the way of training development resources, however. Requirements for efficient and effective training of all the critical skills will not decrease. Thus, tools such as the training system estimation models may be increasingly important in the wise choice of Army training systems.

FRONT END ANALYSIS AND TRAINING ESTIMATION

Training System Development

In the normal development of training systems, there will always be one constant: never enough time and money to generate thoroughly and evaluate carefully the training system and its components. This fact places enormous pressure to produce training as quickly and as cheaply as possible. In so doing, the "normal" procedure is to use as much of existing training material as possible, to minimize initial design of and design effort for the training system, and to use processes and techniques that are most comfortable to the developers even though they may not be the most appropriate for determining the training solution.

At least three major risks can be incurred if training system development is done too hastily:

1. Not enough attention may be given to carefully analyzing the needs for which the training is designed. It is an unfortunate fact that many personnel and organizational needs are not necessarily best solved by training. Thus, it is important to analyze needs carefully to see if they are amenable to, or appropriate to, training. This initial step has been termed "needs analysis" and may be one of the most omitted steps in training system development (Martelli, 1990; Rummler, 1987). It is not uncommon, for example, to find a lengthy and expensive training course for job tasks that really should be re-designed to unburden the operators and considerably reduce their skill and training requirements.

2. Insufficient time may be given to carefully studying the training target audience. In short, what are the capabilities and limitations of the people who are to receive the training? In practice, training is rarely adjusted to the skills and abilities of the trainees. Part of the problem is that it is not easy to describe satisfactorily the appropriate dimensions for training in the individuals of the target audience. Another problem is determining the best source for information. For example, measurement of the target audience is expensive, while using subject matter experts has been a somewhat less than perfect source of information about the trainees or the tasks for which they should be trained (Harless, 1989).

3. Test and evaluation of the training system may be omitted. Then the question arises, does the training system really train what it is supposed to train? Unfortunately, many training packages may appear to be good but upon test or experience, may neither train what is expected nor adequately train what is required. It is possible to get excellent training efficiency (i.e, learning demonstrated) without training effectiveness (i.e., skills learned

and transferred to job performance). Unfortunately, also, the lower limit is not zero; it is possible to inadvertently achieve negative transfer of the job tasks. This result is worse than no training at all.

Front End Analysis

Many of these problems can be avoided or minimized if careful attention is given to the design of the training system before a commitment to actual hardware and software is made. In general, this approach requires a systematic process of development. As Gropper and Ross (1987, page 196) put it:

> Consistently effective training outcomes are attainable only if training professionals have at their disposal a training development process that is systematic, generalizable, and valid.

Such a process is discussed in the next section.

As a part of a general process, much depends upon "front end analysis" and the care taken to define a good training system. Seidel and Wagner (1980) gave the following definition:

> Front end analysis (FEA) is a process that evaluates requirements for manpower, personnel, and training (MP&T) during the early stages of the military system's acquisition cycle. Its purpose is to (1) determine manpower, personnel, and training (MP&T) requirements under alternative system concepts and designs, and (2) estimate the impact of these MP&T requirements on system effectiveness and life cycle costs. Its end product should be the information needed to assume that effective resources (human, equipment, materiel) will be available when and as required for each system to achieve its intended contribution to military readiness and effectiveness. (page 1)

The key elements of front end analysis for training are as follow. First, training requirements have to be established. Second, alternate <u>conceptual</u> solutions should be considered and evaluated. Third, the impact of the requirements and the solutions on cost should be predicted. Fourth, front end analysis provides the <u>information</u> base from which actual training systems will be designed, developed, and deployed.

Front end analysis therefore places the emphasis on initial conceptual analysis and evaluation before any commitment to any real training system design. Its goal is to suggest the most cost-effective means for solving training problems. Finally, it is a decision-making tool in that it provides data and information from which decisions can be made about the potential effectiveness of a training system before substantial investments are made in developing and fielding the system.

Predicting Training System Performance

Therefore, for front end analysis to be successful, it must have a means for predicting training system performance based on conceptual rather than actual designs of the training system. Further, the predictions should be quantitative in nature if possible. That is, actual predictions of relative cost effectiveness in dollars should be made. Other predictions that may be needed include training efficiency and training effectiveness. The decision problem is really twofold. First, there needs to be some realistic idea of how many resources have to be given to the projected training system if it is to work. Second, there needs to be some evaluation of the value of the projected training system to see if it is worthwhile developing and using. Training systems, and particularly training devices and simulators, have become very expensive, and it is reasonable that judgments be made about the predicted value of the system or device before large amounts of money are spent on them.

Need for Training System Estimation Models

To make these kinds of predictions in quantitative terms to enable decision makers to allocate training resources, training system estimation models must be used. These models will provide quantitative performance and cost estimates about projected training systems. At least five general requirements can be stated for such models:

1. The models must be reasonably adequate representations of the training process and system. They must contain the major elements, dimensions, and processes that the training system would have. They must be sensitive to all parameters that may be expected to influence training significantly.

2. They must predict training efficiency quantitatively. One of the most important predictions of this type is the shape of the learning curve, which will then dictate predictions of what and how much will be learned, and quantitative estimates of how long it will take (Lane 1987; Schneider, 1989). One of the most important predictions that the model should try to make is how long training will last--probably in terms of course hours.

3. They must predict training transfer. We train so that there will be better job performance and positive transfer of the training to the job, thus maximizing job performance. The models must then predict quantitatively the degree of transfer of training. Further, since the possibility of negative transfer always exists, that must be predicted as well.

4. The quantitative predictions must be reliable and credible. It is not too difficult to build a model that will generate numbers; the question will be, are they reliable and valid? In short, are the numbers repeatable and do they predict actual job performance?

5. The models must be usable by many different kinds of people (e.g., training system designers, decision makers, operational users) who have a tremendously different background for model use and acceptance. Some critical participants reject any sort of mathematical model even though model successes have been demonstrated for centuries in science and technology and have been widely used for military operations (Hughes, 1984). Others seem to accept models quicker than they do reality.

These five criteria are idealistic. All models will fall short of these requirements to some extent. For example, most real-life processes including training are extremely complex, and their representation in a model will always involve simplifications. Some models may be able to predict training efficiency but not training effectiveness. No mathematical model in any area of science and technology has yet proved to be exhaustively correct, but many have shown enormous practical utility. The heart of the discussion that follows is concerned with specific training estimation models.

TO MODEL WHAT?

The Systems Approach to Training

Previous reference has been made to the comment of Gropper and Ross (1987, page 196) that effective training requires a systematic training development process. Such a process does exist under the names of the systems approach to training (SAT) or instructional systems development (ISD). Both SAT and ISD began to emerge about 25 years ago (cf., Kaufman, Corrigan, & Nunnelly, 1966) in response to a widespread feeling in the training community that training development was inadequate, incomplete, and uncontrolled. Further, there was a perceived need for the process of training system development to consider a series of steps from initial concept to fielded training system to ensure that all critical features of the future training system were accounted for.

Within the military training community, there was a rather rapid response to the concepts of SAT and ISD. By 1970 (Department of the Air Force, 1970), the Air Force had institutionalized an instructional system development procedure. Similar steps were taken by the Army and the Navy (e.g., Department of the Army, 1975).

There are several different versions of the ISD methodology. One reasonably complete summary has been given by Meister (1985) based on several published versions of ISD. This summary is shown in Table 1 and assumes five basic functions in developing and fielding a training ystem: analyze, design, develop, implement, and control training. The importance of the ISD methodology here is to illustrate where training estimation models would be used in the ISD process and to give some more detailed notions of what would be needed in the training estimation model if it is to be usable.

To show the kinds of variations the ISD model can assume, a second version is shown in Table 2. In this version, there are five basic functions: define training needs, state training objectives, specify management and delivery plan, develop courses, and implement the management and delivery plan. The differences between the two models are principally in emphasis. Of particular note here is that each starts with an analysis of the job and the tasks the human must perform. It is assumed that the foundation of all steps in training system development is a thorough statement of the human tasks for which training is being developed. Following that, there will be differences in the importance given to such critical items as statement of training objectives, measurement of training performance, and evaluation and validation of operational training.

Since the 1970s, some version of the ISD methodology has permeated all military and civilian training. The goal has been to produce training that is more complete and more effective. Meister (1985, page 184) points out, "Like other design tools, ISD can really be evaluated only on the basis of <u>utility</u>: Does it work and work well?" It is really very difficult to answer that question empirically since it would require some kind of long-range comparative evaluation of two or more training systems designed for the same training problem using and not using the ISD procedures. It is doubtful that resources will ever be provided for that kind of methodological verification. In the meantime, however, the notions seem particularly useful in reminding all training system developers of the many items and dimensions that have to be accounted for in the development of a good training system.

Table 1

Stages in ISD Development
(Meister, 1985)

1.0	ANALYZE TRAINING
1.1	Analyze job
1.2	Select tasks for training
1.3	Construct job performance measures
1.4	Analyze existing courses
1.5	Select instructional setting
2.0	DESIGN TRAINING
2.1	Develop objectives
2.2	Develop tests
2.3	Describe entry behavior
2.4	Determine sequence and structure
3.0	DEVELOP TRAINING
3.1	Specify learning events and activities
3.2	Specify instructional methods
3.3	Review select existing materials
3.4	Develop instruction
3.5	Validate instruction
4.0	IMPLEMENT TRAINING
4.1	Implement instructional management plan
4.2	Conduct instruction
5.0	CONTROL TRAINING
5.1	Conduct internal evaluation
5.2	Conduct external evaluation
5.3	Revise system

There have been some attempts to "measure" the utility of the ISD process, principally in terms of how it has been used and what benefits were perceived to have been gained from it. Because of the methodological work in the 1960s, many procedural variations of the ISD process were available (cf. Montemerlo & Tennyson, 1976), and by the end of the 1970s, there were attempts to evaluate how well ISD was doing (cf. Vineberg & Joyner, 1980). That widespread use of the method was not universal is perhaps not surprising; the method is lengthy, and some of the procedures needed to implement the steps listed in Tables 1 and 2 have not been specifically established. Even another decade later, we still do not know how to do some of these steps precisely and certainly not as as a standardized procedural technique. There remains, for example, considerable disagreement as to what procedure for task analysis (Table 1, Step 1.1; Table 2, Step 1.1) is best for training system development.

There is another way, however, in which conceptual frameworks such as ISD can be helpful in improving practice. The actual steps and procedures may not be needed so much as the basic concepts and ideas. One can therefore see the fundamental impact of the SAT and ISD in the Army's approaches to training as well as on the procedures used for training materials development (Department of the Army, 1975, 1988a, 1988b, 1990). The primary document establishing the principles of Army training stresses consecutively: training goals, mission-essential task list (METL) development, planning of training, execution of training, and assessment of training (Department of the Army, 1988a). This is very much in the spirit of ISD.

Table 2

A Second Version of the ISD Process

1.0	DEFINE TRAINING NEEDS
1.1	Analyze job tasks
1.2	State required capabilities
1.3	Describe entry qualifications
1.4	Identify capability and qualification differences
2.0	STATE TRAINING OBJECTIVES
2.1	Determine job performance standards
2.2	Describe behaviors, conditions, and standards
2.3	Identify specific knowledge requirements
2.4	Identify specific skill requirements
3.0	SPECIFY MANAGEMENT AND DELIVERY PLAN
3.1	Sequence training objectives
3.2	Specify learning activities
3.3	Identify instructional delivery needs
3.4	Identify training device needs
4.0	DEVELOP COURSES
4.1	Review and select existing materials and devices
4.2	Develop additional materials and devices
4.3	Develop tests and practice exercises
4.4	Verify all training objectives are addressed
5.0	IMPLEMENT MANAGEMENT AND DELIVERY PLAN
5.1	Conduct courses
5.2	Validate training
5.3	Revise training

Training Estimation Models and ISD

The basic question of this chapter is, where do training estimation models fit in the ISD procedure? The answer is that the estimation models are principally for evaluating training system designs; the ISD procedures are used to establish the basis for creating the training system designs that are evaluated. In general, the training estimation model is a supportive design tool that allows for prediction of possible training success before the actual development of training systems. They allow for comparisons of alternate training system design concepts that might be developed by the ISD process.

Thus, the ISD process as shown in Tables 1 and 2 must be seen in the context of training system design. It is assumed that there will be conceptual designs of alternate systems for a given training problem. It is also assumed that before actual commitment to development, there needs to be an evaluation step to select the most favorable option for development. That judgment could be made in many ways, such as the quality of training alone, but it is probable that some cost-effective and quantitative predictions will be required. Presumably, the option selected for development offers the probability of the most effective and least expensive of the alternatives available for training. Parenthetically, it is more probable that the final selection will be some "best" combination of training effectiveness and cost.

Some General Criteria

Given this discussion of the ISD process, it is possible to expand the five general requirements for a training estimation model listed in the preceding section. These additional criteria include the following 12:

1. The model must be available during conceptual design. The model is fundamentally a tool for predicting best options before development, although it may also be useful in assisting design.

2. The model must represent the tasks or skills to be trained. These tasks and skills are the basic materials for all training system predictions and developments.

3. The model must incorporate specific training objectives and standards. In general, the models cannot provide these objectives and standards; they must be specified independently of the training system itself.

4. The model must be responsive to anticipated student characteristics. There will be a need to know what abilities and skills students bring to the training situation since such abilities and skills can complement training.

5. The model must provide comparisons with existing training courses and materials. Very rarely are training courses developed outside a chain of pre-existing training conditions; it is important to see what previously developed training materials can be used and how the performance of the new training system compares with that of the old.

6. The model must be sensitive to fundamental variables in human learning such as massed versus distributed practice, the role of feedback in learning, and conditions that enhance the possibility of positive transfer.

7. The model must present alternate training strategies and conditions of training. For example, there is always a question of where training should take place: in the classroom or on the job. The model should be able to predict the consequences of these kinds of differences.

8. The model must provide quantitative predictions in data terms that might be measurable. The ISD process calls for the development of good measures of training and job performance measurement (Table 1, Step 2.2; Table 2, Step 4.3). In so far as the model predictions are commensurable with the data derived from the use of these measures, there can be a validation of both the model and the training process.

9. The model must compare alternate training media and training devices options. This is the era of high technology training, and the use of training media and devices is the rule rather than the exception in current training programs. A very important goal for the model will be to predict the best fit between media and training objectives.

10. The model must reflect training support and training management requirements and objectives. The delivery of training is a significant cost

factor in all training systems, and cost-effective predictions cannot omit those features associated with training logistics.

11. The model must reflect the potential role of instructors and training support personnel.

12. As stated above, the model must predict training efficiency and training effectiveness as a combination of all the training system variables.

No model will be able to fulfill all these criteria. However, they can serve as indicators to evaluate the models in comparison to each other.

QUANTITATIVE ESTIMATION MODELS

A Short History

Attempts to construct quantitative prediction models for training systems have been going on for the past 20 years. Table 3 shows 36 of the models that have emerged during that period of time. The dates given in Table 3 represent times when the models were available for use. Table 4 cites sources for the models listed in Table 3.

From the beginning, the focus was on the development of cost-effectiveness models. That is, researchers were trying to develop quantitative models that would predict the cost effectiveness of new training systems. Almost all the models listed in Table 3 have cost-prediction routines.

Some general characteristics of the models include

1. Attempts were made for all the models to have immediate system applications. In some cases (e.g., B1-SAT and LAMPS) the models grew from specific weapon systems development programs.

2. Most of the models were developed independently with little apparent influence from and dependence on previous models. There were, however, noticeable exceptions in five cases in which models were developed sequentially:

TECEP--TRAINVICE--DEFT
TECEP--TEEM--TCA--TDDA--TDDSS
CHRT--TRAMOD
CHRT--HARDMAN I--HARDMAN II (MIST)
AIDS--AIMS

Perhaps the longest development period has been that associated with HARDMAN, first developed by the Navy in the 1970s, one version of which was then adopted by the Army. HARDMAN III is being developed, but the HARDMAN tools are intended for the entire manpower, personnel, and training community, and the training models are just part of much larger model programs.

3. Most of the models are computer based. Probably none of them was simple enough to be "manual." All the models require large data bases and the handling of those data bases. All the models have relatively complex algorithms, some of which would be impossible to solve manually in reasonably finite time.

4. There has been a heavy emphasis on media selection. Some of the models are built primarily to focus on media selection (including training device selection). This is perhaps a natural development, considering the tremendous advance in media technology for the past 15 to 20 years. The subproblem of media selection has perhaps been far over-emphasized in light of the many variables that influence learning in a training system.

Table 3

A Summary of Training Estimation Models: 1972-1990

Date	Acronymn	Title
1972	TAEG	Training analysis & effectiveness group models
1974	B1-SAT	B-1 systems approach to training
1974	TECEP	Training effectiveness cost effectiveness prediction
1975	CARAF	Combined arms research and analysis facility
1976	CHRT	Coordination of human resources technology
1976	TRAINVICE	Training device effectiveness model
1976	TEE	Training effectiveness evaluation
1977	MODIA	Method of designing instructional alternatives model
1977	MIL-T	Military specifications & training requirements for aviation weapon systems
1978	ASSET	Acquisition of supportable systems evaluation technology
1978	TCA	Training consonance analysis
1978	TEEM	Training effectiveness estimation model
1978	MDMA	McDonnell-Douglas media allocation model
1978	AIDS	Automated instructional development system
1978	TDDA	Training developers decision aid
1978	TRAMOD	Training requirements analysis model
1979	LAMPS	LAMPS Mark III methods/media selection model
1980	HARDMAN I	Hardware and manpower comparability analysis methodology I
1981	ATM	Analogous task method (training cost model)
1981	CIEA	Cost and information effectiveness analysis
1981	DORAC	Device operational assessment capability
1981	LBM	Learning-based model for media selection
1982	TDDSS	Training developers support system
1983	CASDAT	Computer-aided system for developing aircrew training
1984	DEFT	Device effectiveness forecasting technique
1985	FORTE	Forecasting training effectiveness
1985	HARDMAN II (MIST)	HARDMAN II (man-integrated systems technology)
1985	CBP	Comparison-based prediction
1987	---	Isoperformance
1987	MATS	Model aircrew training system
1987	TASCS	Training analysis support computer system
1987	TDS	Training decision system
1987	TECIT	Training effectiveness and cost effectiveness interactive technique
1988	AIMS	Automated instructional media selection
1988	OSBATS	Optimization of simulation-based training systems
1990	---	Model for skill-based training

Table 4

Documentary Sources for Training Estimation Models

Acronymn	Source
TAEG	Brady (1973)
B1-SAT	Mitchell and Ranney (1974)
TECEP	Braby, Henry, and Morris (1974)
CARAF	Informal technical reports by Vector Research
CHRT	Golclowski, King, Ronco, and Askren (1978)
TRAINVICE	Goldberg and Khattri (1987)
TEE	Carter (1982)
MODIA	Carpenter-Huffman (1977)
MIL-T	Kribs, Simpson, and Marks (1983)
ASSET	Predis (1984)
TCA	Hawley and Frederickson (1983)
TEEM	Jorgensen and Hoffer (1978)
MDMA	Various internal McDonnell-Douglas technical reports
AIDS	Kribs, Simpson, and Marks (1983)
TDDA	Pieper, Guard, Michael, and Kordek (1978)
TRAMOD	Predis (1984)
LAMPS	Kribs, Simpson, and Marks (1983)
HARDMAN I	Various DRC technical reports; Zimmerman, Butler, Gray, and Rosenberg (1984)
ATM	Jorgensen and Hoffer (1978)
CIEA	Hawley, Brett, and Chapman (1982)
DORAC	Hawley and Dawdy (1981)
LBM	Gagne, Reiser, and Larsen (1981)
TDDSS	Hawley and Fredericksen (1983)
CASDAT	Marcue, Blaiwes, and Bird (1983)
DEFT	Goldberg and Khattri (1987)
FORTE	Fishburne and Rolnick (1985); Pfeiffer, Evans, and Ford (1985)
HARDMAN II (MIST)	Mannle, Guptill, and Risser, D.T. (1985)
CBP	Klein, Johns, Perez, and Mirabella, (1986)
Isoperformance	Jones, Kennedy, Kuntz, and Baltzley (1987)
MATS	Fishburne, Williams, Chott, and Spears (1987)
TACS	Butler and Gebhard (1987)
TDS	Collins, Hernandez, Ruck, Vaughan, Mitchell, and Reuter (1987)
TECIT	Adams and Rayhawk (1987b)
AIMS	Kribs, Simpson, and Marks (1983)
OSBATS	Sticha, Blacksten, Knerr, Morrison, and Cross (1986)
Skill-Based	Sticha, Schlager, Buede, Epstein, and Blacksten (1990)

5. All are complex models relative to modern operations analysis and modeling (Hughes, 1984). Most are deterministic, although some of the later models are stochastic (e.g., optimization of simulation-based training systems (OSBATS)), but all are multi-variable.

6. Few have survived to be institutionalized or to achieve sustained usage beyond the research stage. The exceptions appear to be HARDMAN (hardware and manpower), CBP (comparison-based prediction), LBM (learning-

based model for media selection), and AIMS (automated instructional media selection) as an outgrowth of AIDS (automated instructional development system). Why some models have persisted while others have not is really not clear.

7. Most of the models appeared to receive inadequate development and test time. There might have been a consistent underestimation of the challenge and complexity that each model builder was facing, or there simply nay not have been sufficient resources. In one sense, most of the models were "incomplete" in that the researchers were well aware of additional refinement and expansion that would have been desirable for the model.

The Problem of Documentation

Adequate and complete public documentation is available for very few of the models. It would appear that time was usually inadequate to document the model system and particularly the computer code for the software algorithms. It would therefore be impossible to reconstruct some of the earlier models on the list in Table 3.

Second, rarely was archival documentation handled appropriately. Many of these models appeared in organization technical reports, which were not deposited in any of the national information retrieval systems. In some cases, the reports appear to have disappeared forever. Only some private libraries still hold copies of the original reports.

Third, a few review articles and reports have appeared, trying to describe and evaluate some of the models: Adams and Rayhawk (1987a), Braby (1973), Carter (1982), Goldberg and Khattri (1987), Knerr, Nadler, and Dowell (1984), and Rosen, Berger, and Matleck (1985). However, in no secondary source about training known to us has any mention been made of these models. (Secondary sources include such texts as "Principles of Instructional Design" by Gagne, Briggs, & Wager, 1988.) Thus, the models do not appear to have moved into the general technology base of the training field.

Table 4 is an attempt to find and cite sources for all the models listed in Table 3. In about 30% of the cases, no formal reports or archival documentation has been found. Reference has been made either through the review articles or through private copies of draft technical reports which received no additional publishing processing.

Before the history of this technological area is lost, it might be worthwhile for some organization or professional society to institute an archival store for documentation in this area. Appeals to past authors might well generate some of the "lost" documentation. At the least, it would be desirable to have the history of model development in this area available so that past mistakes would not be repeated, and past successes might be augmented.

EVALUATING TRAINING ESTIMATION MODELS

Criteria for Evaluation

There are many criteria by which we could evaluate the training estimation models listed in Tables 3 and 4. The following five criteria were selected upon which to evaluate the 36 models:

1. Does the model predict cost effectiveness? Questions that concern the value and the cost of the training system are of fundamental importance.

2. Does the model predict training efficiency? There must be some indication about the course of training and how well training is being accomplished. Metrics about this, such as training time, will be essential. While there might be a need for the learning curve, it seems very desirable to predict some estimate of the terminal performance state that results from training. In short, to what degree were the training objectives achieved?

3. Does the model predict training effectiveness? Training exists to improve job performance, and it seems very reasonable to expect predictions about what parts of the job will be improved and how much. This prediction is actually very difficult to make.

4. Does the model predict appropriate media selection? Here the term "media selection" is used in the broadest sense, including, for example, training devices and simulators. In this view, training devices are media as are interactive video discs or sound and slide shows. They are simply different media for presenting training materials.

5. Last, is there sufficient documentation to reconstruct the model? In the last section, much concern was expressed about the loss of documentation that has seemed to accompany these models--if appropriate documentation existed at all.

An Evaluation Matrix

Table 5 presents a summary evaluation of the 36 training estimation models considered in this report in terms of the five criteria. These include four types of predictions and the quality of the documentation for the models.

For the first three types of predictions (cost effectiveness, training efficiency, and training effectiveness), most of the models (22 for each category, although not the same 22 in each) provide some kind of prediction. With respect to media selection, 25 of the 36 models have some form of predicting and evaluating media choices. This predominance is not suprising since media selection has been perhaps the major single concern in the area of the training estimation models other than predictions of cost effectiveness.

Looking at documentation, it is almost as if following 1981 (i.e., following cost and information effectiveness analysis [CIEA]), the quality of documentation improved significantly. Thus, in the last decade we can see that not only has documentation of a more complete and archival nature occurred but also reviews of this model literature have appeared.

A major qualification for this entire literature and these models should be noted. In most of the cases, the models are designed for the training system environment only. There is no concern for variables beyond the training problem. In a few cases, trade-offs beyond training are of interest; for example, the Isoperformance model (Jones et al., 1987) is a model specifically based on the generation of trade-offs between training, personnel selection, and equipment variables to get constant levels of system performance. Beyond these, a few of these models (e.g., HARDMAN I, HARDMAN II-MIST) are parts of much larger manpower, personnel, and training (MPT) models that are trying to predict the entire scope of the use of human resources in system design and operations. In these latter cases, it is not surprising that some variables of great interest in the training-specific models are not included. For example, details of course design are probably not of interest in larger MPT models except as part of the general training cost.

Table 5

A Summary Evaluation of Training Estimation Models

Model	Predictions* CE	TE1	TE2	MS	Documentation	Comment
TAEG	+	+	+	+	Fair	Replaced by TECEP
B1-SAT	+	0	0	+	Poor	Developed for B-1
TECEP	+	+	+	+	Good	Several positive evaluations
CARAF	+	0	+	+	Poor	For non-expert users
CHRT	+	0	0	0	Fair	Used later in ASSET development
TRAINVICE	0	+	+	0	Poor	In 1984 changed to DEFT
TEE	0	0	+	+	Fair	Evaluates alternate training
MODIA	+	+	+	+	Fair	Compares alternate training
MIL-T	0	+	0	+	Poor	Good for refining objectives
ASSET	+	0	0	0	Fair	Full demonstration never completed
TEEM	+	+	+	+	Fair	Replaced by TDDA and TDDSS
TCA	0	+	0	+	Fair	Modification of TEEM
MDMA	0	+	0	+	Poor	Originally used for F/A 18
AIDS	0	+	0	+	Fair	Relatively easy to use
TDDA	+	+	+	0	Fair	Integrated with TDDSS in 1983
TRAMOD	+	+	0	+	Poor	Submodel of ASSET
LAMPS	0	0	0	+	Poor	Quick and easy to use
HARDMAN I	+	0	0	0	Fair	Part of larger MPT model
ATM	+	+	+	+	Poor	Consistent with LCSMM predictions
CIEA	+	0	+	+	Fair	Uses MAUM technology
DORAC	+	0	+	+	Fair	Uses CIEA data
LBM	0	+	0	+	Good	Solid learning theory basis
TDDSS	+	+	+	0	Fair	Apple II based
CASDAT	0	+	0	+	Fair	Never applied
DEFT	0	+	+	0	Fair	Based on TRAINVICE
FORTE	0	0	+	+	Fair	Evaluation method for TEE
HDM/MIST	+	0	+	+	Fair	Part of a larger MPT model
CBP	+	0	+	+	Fair	Strong use of expert opinion
Isoper	0	+	+	+	Fair	Trade-off model for design
MATS	0	0	+	+	Fair	Useful for training management
TASCS	+	+	0	+	Fair	For ISD analysis and design
TDS	+	+	+	0	Fair	Predicts best training allocation
TECIT	+	0	+	0	Poor	Incorporates DEFT and FORTE
AIMS	0	+	0	+	Good	Relatively easy to use
OSBATS	+	+	+	+	Good	Development not complete
Sk-Based	+	+	+	+	Good	No hardware or software implementation

*Predictions:

CE = cost effectiveness
TE1 = training efficiency
E2 = training effectiveness
MS = media selection

+ = YES
0 = NO

Considering that the models listed in Table 5 have been generated basically in the last 20 years, it is remarkable that so many models of such great complexity and intended scope have emerged. These models represent an attempt to place training on a sound quantitative estimation technology basis. When decision makers have demanded better estimates of what training and training costs are best for variable amounts of resources, these models have been an attempt to respond with great sophistication. That they are mostly incomplete and not widely used as yet may be more a reflection of the very limited amount of time they have existed. It is really too early to "evaluate" these models in terms of their potential; they probably need another 20 years to evolve and mature, assuming that resource commitments will be made to them.

CONSTRAINTS IN USING TRAINING ESTIMATION MODELS

When Not to Use the Models

There are surely many training system design and development situations in which the training estimation models should not be used. At least four such situations can be noted:

1. The training system cannot be adequately modeled. This would mean that the real training system is so complex, has so many elements, and will be used by so many people with large numbers of media and resources that the parts and processes of the system defy modeling description. This might particularly be true if a standard form model were being used, which could not be easily expanded for greater complexity. The model will always work at least in part as a function of the adequacy of the description of the process being modeled. To the extent that parts of the training system cannot be modeled, the predictions of the model will suffer to varying degrees.

2. Training data for the model are excessively incomplete. Many kinds of data must be generated to make a training estimation model work, such as cost, learning curves, and transfer of training data. Despite many statements to the contrary in the literature, cost data can be adequately obtained although not easily. Technical cost data and accounting systems can generate satisfactory cost data. Not all cost data cells will be complete, however, and not all will be accurate, but most of the cost of training can be estimated within very reasonable limits.

What is very difficult to obtain are data about training and transfer performance. While the general shape of the learning curve may be a power function (cf. Lane, 1987), there sometimes seems to be as many separate power functions as there are learning curves. While learning tends, perhaps, to be negatively accelerated, there are many exceptions. What we may need to know is how the training sample will learn given tasks, and we may need to know the exact functions. These data are rarely available. Part of the problem has been that specific learning curve data have not been accumulated so that we can have some idea of the classes of learning curves that might be associated with different kinds of human tasks.

3. Setup time may be too limited. No matter how general and flexible the training estimation model may be, it takes time to set up the model, test it for the particular application, and apply it to the specific problem. Many training programs seem to have to be initiated in very short time spans with

little time allotted for careful analysis and design. Very frequently, the time limits are such that the first step is a frantic search for already available training packages.

Part of the solution depends on how frequently a training system estimation model installation will be used. When only a few training systems are developed and they are used infrequently, it is difficult to justify the cost of installing and maintaining a model that could be constantly up and ready for use. This is an investment cost that not many organizations are willing to pay, and it is particularly a problem if the estimation model is complex, requiring extensive computer support. A larger cost for such a setting is for the people who must run and maintain it. If, however, this investment is made, then short setup times can be accomplished because the model is ready for quick use.

4. But the biggest single problem in deciding whether to use training estimation models comes from the total cost of the models. Some of the models that have been described here would cost approximately $100,000 to set up and maintain yearly. The question is, where do these costs get distributed into the overall cost of the training system? Are these costs "cost effective?"

The "Return on Investment" Argument

For training estimation models to be justified, it is commonly felt that they should provide increments in the training system that would not have been obtained without their use and that a substantial savings in total training system cost should be achieved. Training estimation models, therefore, must be able to show a "return on investment" when they are used in training system design and development.

Perhaps it is useful to distinguish between two training situations. The first occurs when there is no prior training system, and the task is to design and develop a new training system. The second situation occurs when a training system is in place and the task is to design and develop a replacement or an enhancement.

In the first case, the training estimation model can be very useful if the model can generate equal training benefit alternatives. That is, the model can suggest alternate training solutions and then derive the relative cost benefits of the solutions. This procedure not only has the advantage of putting the selection of a training system on some sort of rational basis but it also allows for a reasonable comparison of actual alternatives. In addition, it provides some sort of rational economic setting to determine what the training should cost. Very frequently, budgeting for training must depend upon guessing since no one can systematically estimate what the new training should cost. Sometimes these estimations can be sobering. In one case, a guess was made that an interactive video disk training package could be designed and developed for $300,000. The training estimation model showed that the cheapest of several video disk alternatives was closer to $2,000,000, and the model suggested that video disk might not be the most appropriate medium for the training. In this kind of case, much money can be saved by avoiding poor training solutions. This is particularly true when very expensive high technology training media choices must be made.

The second situation is when a training system exists and the goal is to supply a new replacement. Here, very probably, the cost of the system in place is known. Using an example developed by Shipp (1981), we can distinguish at least three major sources of training costs: total courses offered, total training hours, and total number of trainees. All these are potential candidates for cost savings in the new replacement training system:

• It may be, for example, that fewer training costs can be incurred if the total number of courses can be reduced. One course may be more effectively taught to encompass the training content of another. One finds, therefore, a constant emphasis on reducing the number of training courses while trying to place the content in a remaining course.

• No single variable in training system cost reduction gets more attention than total training hours. Sometimes it would appear that every vocational training course is being examined to see how much it can be shortened. Surely, there is no question that costs are reduced when the number of training hours are reduced, but one potentially undesirable consequence is that training may now become excessively concentrated (massed rather than distributed). It is well known from the scientific literature about human learning that massed training as compared with distributed training reduces both training effectiveness and skill retention.

• Parenthetically, cost savings programs that concentrate on reducing training hours seldom consider what impact, if any, this has on other major cost items in the training system. If reducing training time has no impact on the care and maintenance of training facilities or devices, it is doubtful if simply reducing training time will result in major savings in the total training system.

• Perhaps the major consideration in training cost reduction is the number of trainees who will use the system. Many training requirements are established for a small number of operators and maintainers when these trainees have a legitimate training need but when there simply are not many students. The use of an expensive training estimation model in design and development will increase radically the cost of training per trainee. It is unfortunate but true that the larger the number of trainees, the more the use of a training estimation model will be feasible from a cost benefit point of view.

However, all this avoids the biggest single problem in true cost-effectiveness estimation and measurement of training systems. What is the value of the training? Here, clearly we need some way of estimating the importance of training to job performance and the subsequent impact on system performance that training can bring. Training at least should enhance personnel skills; these skills in turn should result in better system performance measures such as increased productivity or improved efficiency, which should increase the effectiveness of the organization. It is very difficult to estimate the value of training in these senses. When such measurement has been attempted, the results have been very encouraging. Shipp (1981, page 288), for example, cites an employee assistance program which "...yielded a $1.61 return for each $1 invested." While the quantitative estimation of the value of training is not easy and may be subject to considerable error (cf. Hunter, Schmidt, & Coggin, 1988), that estimation should be attempted. Without such a value estimation, the only rational result of cost effectiveness of training analysis is to eliminate training altogether.

One perhaps very useful notion in establishing training value is that of a "worth of ownership assessment" (Allbee & Semple, 1981). This idea is a specific assessment of the benefits of training: "The general goal of a worth of ownership assessment is to objectively identify the training system whose capabilities best meet training mission requirements" (page 135). This procedure starts with identifying desired training objectives, a step which is frequently ignored in training system development. Allbee and Semple (1981) were able to identify 65 worth factors in terms of eight general areas: political, management and administrative, resource management, operations and tactics, training, personnel, training effectiveness, and training device technology. In this procedure, measures were defined for each of the 65 factors and a process described for calculating worth of ownership.

It may be easier to estimate some system performance correlates of worth of training system ownership in the civilian world, but it is surely not impossible to find such correlates in military training systems, at least as reflected in the famous statement of Marshall Kutsekov: "Train hard, fight easy."

The Cost Data Base Problem

Reference was made to the frequent problem of finding adequate cost data for training estimation models. This issue should be looked at a little closer. First, it is not difficult to find cost structure models for training systems in the military (cf. Knapp & Orlansky, 1983). As might be expected, a reasonably accurate cost model will contain many individual cost elements. Knapp and Orlansky (1983, pages 50-52) divide these elements into three basic cost categories: research and development, initial investment, and operating and support costs.

For a given training system, it is difficult to assess that system for research and development costs that might well be distributed over many different specific training system developments. There is little question, however, that for every training system development, initial investment and operating costs should be examined separately. Kinkade (1980) demonstrated this with data shown in Table 6, based on a set of specific Navy training courses.

Table 6

Relative Cost Differences for Training Media*

Development costs: media	Operating costs: Ratio to lecture	Ratio to lecture
Lecture	1.0	1.0
Workbook	1.8	1.1
Random access slide	2.1	1.1
Random access fiche	2.7	1.4
Sound/slide	3.5	1.4
Videotape cassette	3.5	1.4
Beseler cue-see	12.7	1.2
Programmed text	3.0	1.4
PLATO	3.7	3.1
Random access fiche processor	5.4	1.7
Video disk processor	7.0	1.7
TICITT/VCR	24.5	1.9

* Kinkade (1980)

These data show a number of things. First, they show that development costs for various training media can vary tremendously as a function of the particular training medium selected. In general, training media with interactive branching may be considerably more expensive to develop than media without interactive features. Second, once the development cost is paid, the relative operating costs are fairly close to each other and are not markedly different from the operating costs of lectures (with a single exception).

Third, and most important, the data imply that detailed cost data are available for analyses and comparisons such as these shown in the table. Given development and operating costs, one can estimate total life cycle costs if such factors as the number of trainees completing the course, proportion of course modification, and frequency of course modification are known.

Finally, the data base for any training estimation model will be a complex mixture of objectively quantitative and subjectively quantitative numbers. One can very probably obtain an objectively correct number for instructor pay and allowances and even allocate it appropriately to particular training (Knapp & Orlansky, 1983, page 43), but the worth of ownership probably will have to be generated through subjective scaling techniques (Allbee & Semple, 1981, page 151). Among other issues, this may create some problems concerning the basic reliability and validity of training estimation models.

Reliability and Validity

Most of the training estimation models discussed here are deterministic; they should derive the same quantitative estimates on every trial. A few of the models (e.g., Sticha et al., 1990) contain stochastic distributions within some of the subroutines and necessarily will derive different values upon repeated estimation. In this case, however, the stochastic models will generate probability distributions, which may or may not better estimate uncertainty in the real training world. The presence of subjectively scaled values will surely reduce (probably) some of the reliability estimates about model predictions. On the other hand, done properly, inter-rater estimates of scalar values can have very high reliability values. There will be no general answer to the reliability question; each model will have to be evaluated by itself, but it is difficult to see where any serious reliability problems should occur if the model is being exercised properly, either by hand or by computer.

With respect to validity, however, the issue is quite different. To what extent will the predictions of the training estimation model be true? That is, will the model really predict (reasonably) what subsequently happens in the real world?

There is probably no general way to test the validity of any of these models. They are too complex, and any empirical investigation would require large scale iterations during long periods of time that simply are not feasible practically. Further, the model predictions may compare theoretical alternatives (such as a variety of different training media), but there will be no way of predicting and testing these alternatives since it is very improbable that more than one alternative will ever be developed. The major purpose of the estimation model is to evaluate for selection a number of alternatives only one of which will in practice be developed.

In some cases, such predictions as the number of hours that training might take or the level of learning to be achieved will be testable. Tests of this sort would require measurement of training effectiveness in the subsequent operating training world. Thus, if the model makes learning and transfer predictions, then testing for those predictions is surely possible. Such measurement would be essential to evaluate an ongoing model simulation. Measurement of this sort would be needed to improve model processes and predictions.

The Users

Like all models, the training estimation models exist to be used. The question here is, who would use them? We distinguish among training system designers, decision makers, operational users, and training system researchers.

Training System Designers

The primary users for whom the training system estimation models were developed have been training system designers. It has been felt that such models will assist the designers in the better design of training systems. Yet, there is no solid evidence that the models have ever been used by training system designers. In general, they have not been user friendly; they have been rarely stabilized so that they can operate reliably; and they seem not to be too consistent with the way that training systems actually are designed.

Klein (1987), for example, has suggested

> I am skeptical of some of the standard accounts of the design process. Such approaches are often misleading because they portray the designer as overly analytical and continually performing trade-offs and calculations.
>
> In contrast, it seems to me that much of the strength of an experienced designer comes from his ability to recognize the types of problems encountered, to recognize the typical ways of handling such issues, and to recognize the implications of contextual nuances...The overriding strength of a skilled designer is in knowing what problems are typical and what reactions they call for, along with what problems are unique and what approaches are most promising. (pages 175-176)

If the training estimation models are considered design aids, then Klein would probably state they are inappropriate in that they "aid" a process that the designer is not using.

The actual design process, for better or for worse, is probably far more intuitive and creative than rational and analytical. Designers tend to look for solutions of their current design problems, based in large part on the solutions that have worked before or that appear promising. Given the stage of design, Rouse (1986) suggests that designers may be artists creating solutions or analysts evaluating alternatives. They will need different kinds of information (and perhaps models) for the kind of activity they are doing, or their design activities will have to be changed.

That much more analysis might be desirable is certainly stressed by some (cf. Oneal, 1990) who offer a variety of sophisticated requirements, technical, and operational analyses. However suitable the models may be for several purposes, it is probable that (a) many, if not most, training system designers do not know how to use them and (b) the training system design environment does not contain easily available modeling and analytical tools. While the general theory of such analysis is rather clear, there seems to have been little concrete development of tools for the designer, which are user friendly and readily available in the workplace.

Decision Makers

Probably the user group with the most potential for implementing the outputs of training estimation models are decision makers and particularly those who control resources. The models tend to provide comparison data that allow for relative judgments about at least the costs of training systems. They also provide a framework to evaluate changes and alterations in the training system. The dominant theme among resource managers with respect to training appears to be reduction of training. Some decision makers for military training systems and investments in military systems do not appear to fully understand the value of training unless it is made explicit.

The biggest need for this class of users is the credible prediction of the value of training. There must be clear demonstrations of the value of training to the military establishment. Training system estimation models have failed significantly in this objective. Probably what is needed is some clear set of credible correlations between level of training and military system performance.

Operational Users

Much the same need exists for operational users of training systems, who tend to understand in vague and general ways the values of training but who are unable to articulate those values to decision makers. Further, users are expected to establish training requirements so that the entire training system research and development community can respond to the needs of users. Yet, most users are unable to state clearly what those training needs are in terms that can be translated into training system development programs and training system packages.

One of the most widely used cliches of research and development is the stated need for operational user inputs into the research and development and design process. While in general there is a much greater need for interactive inputs between operational users and designers, this requires the operational users to have some understanding of what technology can and cannot provide for them in terms of operational capability. Many operational users misunderstand the capabilities and limitations of new technology and seem to be either over-enthusiastically supportive or inappropriately negative. There should be better ways of explaining to the user the uses and misuses of high technology while allowing the operational user to express what is needed on the battlefield or in the operational process.

Much of the dialogue between users and developers might fit under the heading of what Klein (1987, page 179) has noted as "ill-defined problems" which, he states, "...require two simultaneous processes: 1) goal clarification, and 2) option development." On the one hand, one has to clarify what the user can really use and needs, and on the other hand, specifically what will help the user reach operational goals.

Training System Researchers

For the past decade, it has seemed that most of the work of training system research and development specialists has been to find and develop training system "solutions" and then go look for a problem to fit the "solutions." An implicit (and sometimes explicit) assumption among training researchers has seemed to be that the more complex and high technology the training "solution" is, the better it is. Unfortunately, during the same decade, measurement of training and transfer effectiveness seems to have diminished considerably so that it is very difficult to say what training "solution" is better in training value than another "solution."

Another unfortunate trend has been the overemphasis on high technology training media as opposed to better use of other parts of the training system. There is, for example, far less interest in more effective use of instructors. One trend during this past decade has been the possibility that human instructor roles may be automated through intelligent tutoring systems (cf. Sleeman & Brown, 1982).

There are few definitive data that show high technology training systems and media to be superior in training and transfer effectiveness to less sophisticated training methods. There is the suggestion that high technology training is task dependent. In some cases, high technology training may be superior for some human operator and maintainer tasks; in other situations, low technology methods may be better. What we need are much more comparative data rather than new unfocused technology.

RESEARCH ISSUES

The "State of the Art"

Looking at Table 5, one is tempted to assess the "state of the art" of training estimation models. What one sees is two decades of exploratory development, which have resulted in a collection of mostly unconnected and uncoordinated quantitative models. To a great extent, this has been a period of trying ideas, some of which have been very complex. In one sense, the accumulated "state of the art" has provided a large learning experience from which many major future advances may be possible. One point of view is that the past two decades were necessary to structure and to begin to understand the problem.

Among the <u>strengths</u> that have been achieved so far, at least five stand out:

1. The purpose of building these models has become increasingly clear, and even if any one of them cannot perform exactly and precisely as desired, what is needed has not been ambiguous. The models are simply trying to assist in judgments about the best kind of training system to build. The greater the investment in that training system, the more important it is to have some sort of process by which the best options can be disclosed.

2. The models have made very clear that precise statements of training objectives can only help in the design and development of any training system. In the past many training systems have been developed for "non-training" problems or problems that could have been solved by means other than training (e.g., redesign of the prime system or the job). Training systems have been

developed that trained, if at all, for something other than what one wished to train. By their very nature, these models demand as clear statements as possible of training objectives and goals.

3. The submodels for media selection have been very effective. In a training world where a multitude of media are available, there must be some rational way of selecting the right media--media that will train "best" and at a reasonable cost. These models have provided a number of such submodels. It is not easy to select the "winning" media selection model, however. At least two or three have achieved semi-institionalized status, that is, they are in use as tools for development (Gagne, Reiser, & Larsen, 1981; Kribs, Simpson, & Marks, 1983). More rational ways of selecting appropriate media are essential in a marketing and acceptance climate where high technology "solutions" may be too attractive--they look nice but they may not train well.

4. The models have given us many ways of making quantitative estimates about the cost effectiveness of training. There is a need to know what kind of return on investment we are getting from training when billions of dollars per year are being invested in training. This work is also part of a larger effort to state the utility of all human resources management. These models offer a large range of cost-effectiveness judgments, about the specific training system, about the training system as a part of larger systems, or about costs over the entire life cycle of the training or larger system. The cost-effectiveness parts of these models parallel and were influenced by the general movement during the past three decades toward more quantitative cost-effectiveness measurement. These models are as sophisticated in the domain of cost effectiveness as any such models anywhere. There is a sense in which they are better; they must address specific parameters of multi-dimensional training systems. At the same time, valid estimation of system cost effectiveness is a very difficult problem (Hunter, Schmidt, & Coggin, 1988); some of these models represent major attempts to solve these problems with real systems.

5. The models encourage systematic thinking about training system development. Too often, training systems are developed hastily on the basis of what was done the last time and what high technology media look attractive. One cannot use these models without thinking carefully about the parameters of the specific training system and then making some systematic attempts to evaluate different options to provide the best possible elements to a training system.

That there are major weaknesses in the models should be apparent by now including at least these five:

1. The models are probably too complex and expensive. They represent attempts to solve too many problems with too great precision. The availability of the computer as a supporting tool has not helped in this; one begins to believe all the potential help the computer could give, but then one finds how difficult it is for the computer to actually give the help. Because of their complexity and expense, the models cannot be easily set up at installations other than the original research setting. There are no standardized disks or programs available for rapid installation and use. Re-creating many of these models would be a very time-consuming and expensive undertaking. The models simply have not been developed to the point of standard use. For many of them, because of inadequate documentation, it might not be easy to get program code.

2. Most of the models require too many data. They need very large quantities of training efficiency, effectiveness, and cost data. Many of these data could be obtained, but the additional requirement of rather frequent update would place an enormous data collection and financial burden on the model installation. The models may also need too many data because the algorithms are too complex and the questions they are designed to answer are unnecessarily detailed and precise.

3. To be usable in design and in decision making, models of this type must give rather rapid answers. It should be possible to set up questions for the model and have answers to the questions within a few days if not within hours. For these models, most of them would require a major research effort to set up a problem and months to provide answers. The answers might be very good, but they will surely be too late. A maximum time response limit needs to be established for these models; they must provide reasonable answers to sensible questions in much less time.

4. The models as they are currently configured do not really aid design. Klein (1987) has presented a persuasive case as to why they do not; the models are principally for evaluating designs rather than for creating designs. Most of engineering design proceeds at a rather rapid pace; alternate design concepts of how to make things work are created rather frequently (and often are eliminated even faster). There are not long pauses while some evaluative model compares the options. Recently, some models in the design areas have been incorporating elementary features of evaluation; for example, if a geometric layout exceeds the specified anthropometric dimensions of the intended users, then there can be an immediate feedback. Something like this capability is going to be needed if the models are ever to be a helping part of the design process.

5. At best, the models have been marginally good in training effectiveness predictions. This reflects a general condition in education and training in that we can predict quite well who will do well in training and even how training will proceed (training efficiency), but we do not predict very well how that training will affect job performance. Training is just one of the many variables that influence job performance, and at best, it probably accounts for less than a majority of the variance in job performance. It is understandable, therefore, why there is so much difficulty in predicting what impact training will have on job performance itself. The actual goal of the model, however, should be to do a good job of predicting that part of the job performance variance for which the training was designed.

Researchable Problems

One of the benefits of the work discussed here is that it has opened a world of researchable problems from the standpoints of both training data and methodology. Six such areas of researchable problems could be mentioned among many others:

1. To get better training effectiveness predictions, we need better transfer of training data and relational functions. Some 25 years ago, there was much interest in developing theories of transfer of training. At the heart of these theories was the assumption that transfer was a result of the relationships between what was learned in training and what was performed on the job. While that in general is reasonable, the details are obscure. More transfer-of-training data need to be collected for training systems, and more attempts to understand what is transferred need to be made. For example, it

is possible that transfer laws differ significantly for verbal and motor learning, the latter showing consistent positive transfer over wide ranges of variations and the former showing consistent negative transfer with very slight changes between training and transfer conditions. If this is so, then the nature of the task itself could be a useful distinction in predicting transfer effectiveness. Hundreds of transfer studies are performed each year, and many are reported in the open literature. At the least, it would be good to see what the studies are telling us about the general nature of transfer as a function of the tasks performed.

2. Because the models must have specific parameters and measures of those parameters, work with the models has tended to improve our understanding of the measures of the learning and transfer process. What to measure is a critical problem in both, and it would be worthwhile to look much more carefully at what measures do predict the learning and transfer process, and what do not. That there is no standardization in measurement for education and training is obvious; that there should be is perhaps not so apparent. The models cannot be ambiguous with respect to measures; they must state exactly what measures are going to be predicted.

3. Some of the steps in the current ISD processes are not clear or at least not as clear as they should be. There is much disagreement about, for example, what constitutes a good "needs" analysis or, more basically, what is the most appropriate form for task analysis. Because the products of these analyses are necessary for the estimation models, the models may help in determining what forms are most useful. Surely, the models can place requirements on the analyses in so far as the analyses must provide inputs to the models. No single form of any of these analyses is probably going to become the "standard," but some clarifications and improvements would be welcome.

4. The models should help get better data. The models should help in stating what measures should be taken and what data should be collected using those measures. This effort could help not only the models but could also help the process of measuring the training and transfer process. The models should also assist in the kinds of details about data collection that are so critical such as how frequently measures should be taken and with what precision they should be measured.

5. It is probably the case that the trend for the future should be toward simpler models that ask, and answer, simpler questions. At the very least, four-digit precision is not needed in the predictions for any known questions. Principally, what is asked for is relative rankings with respect to options for training. This can often be obtained with ordinal, or at most, interval scaling. To those rankings can be added cost dimensions and comparisons. If the ranks remain invariant across all scale types, the ordinal level is probably satisfactory. But, what is needed is a more careful look at precisely what questions the training estimation model is supposed to answer and this, in turn, requires one to look closely at the intended user. What does the user need to know?

6. If these models are to be design aids, they must be made more "user friendly," and they must be adapted to the user in the design setting. Perhaps the best way of doing this is to create demonstration applications of the tool where it can be used by several designers in many designs of training systems.

Lessons Learned

During the past 20 years, at least some lessons learned can be suggested with respect to training estimation models:

1. Complex, computer-based, training estimation models take a lot of time, money, and skill to develop.

2. Development without documentation is a waste of resources.

3. There is no substitute for thorough and careful checkout and test. Some of the model problems were attributable to incomplete development, test, and revision.

4. The "value" of training is still not fully understood in many quarters, and cost-effectiveness estimations that stress cost and minimize training effectiveness do not help that uncertainty.

5. These models have demonstrated that one really can quantitatively predict and evaluate alternate options for training before a system is built. A task that remains is to convince the training community that some of the models do work and that they can be helpful in the development of training.

6. Despite the enormous amount of effort and resources that have been put into the development of training estimation models during the past 25 years, very few of the models have received wide acceptance and use. Those people who constitute the training system development community are, for the most part, not comfortable with such models and their complexities. Many have not seen the value of the models or have simply not had the time to explore their use. There has been considerable resistance to change in the training development community and to using training estimation models.

To be fair, that resistance in many cases has been warranted. The models often have been too complex, required too much and sometimes unobtainable data, and demanded high levels of modeling and computer manipulation skills. Further, the authors are not aware of any demonstration of using the models with actual alternate training courses. However, a very substantial technology base of training estimation models has accumulated, and it seems important that this base not be lost or ignored. Training estimation models still hold great promise for the future.

REFERENCES

Adams, A.V., & Rayhawk, M. (1987a). A review of cost and training effectiveness analysis (CTEA): Volume 2. Cost analysis (RN 87-59). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Adams, A.V., & Rayhawk, M. (1987b). Training effectiveness and cost-iterative techniques: Volume 2. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Allbee, K.E., & Semple, C.A. (1981). Aircrew training devices: Life cycle costs and worth of ownership (AFHRL-TR-80-34). Wright-Patterson AFB, OH: Air Force Human Resources Laboratory.

Braby, R. (1973). An evaluation of ten techniques for choosing instructional media (TAEG Report No. 8). Orlando, FL: Training Analysis and Evaluation Group.

Braby, R., Henry, J.M., & Morris, C.L. Jr. (1974). A technique for choosing cost-effective instructional media. Orlando, FL: Training Analysis and Evaluation Group.

Butler, G.J., & Gebhard, E.P. (1987). TASCS III, the computer-based tool that makes the ISD process practical. Proceedings of Technology and Innovations in Training and Education Conference, 339-345.

Carpenter-Huffman, P. (1977). MODIA: Volume 2. Options for course design (AF Report No. R-1701-AF). Santa Monica, CA: The Rand Corporation.

Carter, R.J. (1982). Methodologies for evaluating training products and processes. Proceedings of 26th Annual HFS Meeting, 258.

Collins, D.L., Hernandez, J.M., Ruck, H.W., Vaughan, D.S., Mitchell, J.L., & Reuter, F.H. (1987). Training decision system: Overview, design, and data requirements (AFHRL-TP-87-25). Brooks AFB, TX: Air Force Human Resources Laboratory.

Department of the Air Force (1970). Instructional system development (Air Force Manual 50-2).

Department of the Army (1975). Instructional systems development (TRADOC Pamphlet 350-30, five volumes). Fort Monroe, VA: U.S. Army Training and Doctrine Command.

Department of the Army (1988a). Training the force (Field Manual No. 25-100). Washington, D.C.: Headquarters, Department of the Army.

Department of the Army (1988b). Training: Systems approach to training (TRADOC Regulation 350-7). Fort Monroe, VA: U.S. Army Training and Doctrine Command.

Department of the Army (Draft, 1990). Systems approach to training - Analysis (TRADOC Pamphlet 351-13). Fort Monroe, VA: U.S. Army Training and Doctrine Command.

Fishburne, R.P., & Rolnick, S.J. (1985). Guidelines for conducting a training effectiveness analysis (TEE): Volume 1. The evaluator's handbook (ARI Research Memorandum 85-14). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Fishburne, R.P., Williams, K.R., Chott, J.A., & Spears, W.D. (1987). Design specifications development for the C-130 model aircrew training system: Phase I report (AFHRL-TR-86-44). Williams AFB, AZ: Air Force Human Resources Laboratory.

Gagne, R.M., Briggs, L.J., & Wager, W.W. (1988). Principles of instructional design (3rd ed.). Fort Worth, TX: Holt, Rinehart, and Winston, Inc.

Gagne, R.M., Reiser, R.A., & Larsen, J. (1981). A learning-based model for media selection (ARI Report No. 81-25a). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Golclowski, J.C., King, G.F., Ronco, P.G., & Askren, W. (1978). Integration and application of human resource technologies in weapon system design: Processes for the coordinated application of five human resource technologies (AFHRL-TR-78-6-11). Brooks AFB, TX: Air Force Human Resources Laboratory.

Goldberg, I., & Khattri, N. (1987). A review of models of cost and training effectiveness analysis (ARI RN 87-58). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Gropper, G.L., & Ross, P.A. (1987). Instructional design. In R.L. Craig (Ed.), Training and development handbook (pp. 195-216) (3rd ed.). New York: McGraw-Hill.

Harless, J. (1989). Wasted behavior: A confession. Training, May, 35-38.

Hawley, J.K., Brett, B.E., & Chapman, W.A. (1982). Cost and information effectiveness analysis: Volumes 1 and 2. An improved methodology. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Hawley, J.K., & Dawdy, E.D. (1981). Cost and information effectiveness analysis (CIEA): A methodology for evaluating a training device operational readiness assessment capability (DORAC). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Hawley, J.K., & Frederickson, E.W. (1983). The training development decision support system: An aid to the design of cost-effective training programs. Proceedings of the 27th Annual HFS Meeting, 397-401.

Hughes, W.P. Jr. (Ed.) (1984). Military modeling. Alexandria, VA: The Military Operations Research Society.

Hunter, J.E., Schmidt, F.L., & Coggin, T.D. (1988) Problems and pitfalls in using capital budgeting and financial accounting techniques in assessing the utility of personnel programs. Journal of Applied Psychology, 73(3) 522-528.

Jones, M.B., Kennedy, R.S., Kuntz, L.A., & Baltzley, D.R. (1987). Isoperformance: Trading off selection, training and equipment variations to maintain the same level of system performance. Proceedings of 31st Annual HFS Meeting, 634-637.

Jorgensen, C.C., & Hoffer, P.L. (1978). Predictions of training programs for use in cost and training effectiveness analysis. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Kaufman, R.A., Corrigan, R.E., & Nunnelly, C.I. (1966). The instructional system approach to training. Human Factors, 8(2), 157-162.

Kinkade, R.G. (1980). Comparative cost indexes for the selection of instructional media (NTEC Contract N61339-80-C-0098 Final Report). San Diego, CA: Essex Corporation.

Klein, G. (1987). Analytical versus representational approaches to design decision making. In W.B. Rouse and K.R. Boff (Eds.), System design (pp. 175-186). New York: North-Holland.

Klein, G., Johns, P.G., Perez, R.S., & Mirabella, A. (1986). Comparison-based prediction of cost and effectiveness of training devices: A guide. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Knerr, C.M., Nadler, L.B., & Dowell, S.K. (1984). Training transfer and effectiveness models (HumRRO FR-TRD(VA)-84-1). Alexandria, VA: Human Resources Research Organization.

Knapp, M.J., & Orlansky, J. (1983). A cost element structure for defense training (IDA Paper P-1709). Alexandria, VA: Institute for Defense Analysis.

Kribs, H.D., Simpson, A.C., & Mark, L.J. (1983). Automated instructional media selection (AIMS) (NAVTRAEQUIPCENT Report No. 79-C-0104-1). Orlando, FL: Naval Training Equipment Center.

Lane, N.E. (1987). Skill acquisition rates and patterns: Issues and training implications. New York: Springer-Verlag.

Mannle, T.E., Guptill, R.V., & Risser, D.T. (1985). HARDMAN comparability analysis methodology guide: Volume 3. Requirements analysis. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Marcue, N.C., Blaiwes, A.S., & Bird, R.G. (1983). Computer-aided system for developing aircrew training (CASDAT) (NAVTRAEQUIPMENT Report No. 79-C-0076-1). Orlando, FL: Naval Training Equipment Center.

Martelli, J.T. (1990). TechnicaT training: The analysis phase. Technical & Skills Training, Aug/Sept, 9-13.

Meister, D. (1985). Behavioral analysis and measurement methods. Chapter Six: Training Analysis. New York: John Wiley & Sons.

Mitchell, J.F., & Ranney, T.A. (1974). B-1 systems approach to training task analysis ratings (B-1 Technical Memo No. SAT-7). Wright-Patterson AFB, OH: Air Force Systems Command, B-1 Project Office.

Montemerlo, M.D., & Tennyson, M.E. (1976). Instructional system development: Conceptual analysis and comprehensive bibliography (NAVTRAEQUIPCENT Report IH-257). Orlando, FL: Naval Equipment Training Center.

Oneal, J.A. (1990). Integration of training systems and analyses. In H.R. Booher (Ed.) MANPRINT: An approach to systems integration. New York: Van Nostrand Reinhold. Pages 391-416.

Pfeiffer, M.G., Evans, R.M., & Ford, L.H. (1985). Modeling field evaluations of aviation trainers. Orlando, FL: Naval Training Equipment Center.

Pieper, W.J., Guard, N.R., Michael, W.T., & Kardek, R.S. (1978). Training developers decision aid for optimizing performance-based training in machine ascendent MOS. Valencia, PA: Applied Science Associates.

Predis, R.J. (1984). Computerized technology package--ASSET: Test and evaluation. Brooks AFB, TX: Air Force Human Resources Laboratory.

Rosen, M.H., Berger, D.C., & Matleck, R.K. (1985). A review of models for cost and training effectiveness analysis. Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Rouse, W.B. (1986). On the value of information in system design: A framework for understanding and aiding designers. Information Processing and Management, 22(2), 217-228.

Rummler, G.A. (1987). Determining needs. In R.J. Craig (Ed.) Training and development handbook (pp. 217-247) (3rd ed.). New York: McGraw-Hill.

Schneider, W. (1989). Training models to estimate training costs for new systems. In J.I. Elkind, S.K. Card, J. Hochberg, and B. Messick (Eds.) Human performance models for computer-aided engineering (pp. 215-232. Washington, D.C.: National Academy Press.

Seidel, R.J., & Wagner, H. (1980). Front end analysis to aid emerging training systems (HumRRO Special Report 80-3). Alexandria, VA: Human Resources Research Organization.

Shipp, T. (1981). Calculating the cost effectiveness of training. In R. Zemke, L. Standke, and P. Jones (Eds.) Designing and delivering cost-effectiveness--and measuring the results. Minneapolis, MN: TRAINING Books, pp. 287-288.

Sleeman, D., & Brown, J.S. (Eds.) (1982) Intelligent tutoring systems. San Diego: Academic Press.

Sticha, P.J., Blacksten, H.R., Knerr, C.M., Morrison, J.E., & Cross, K.D. (1986). Optimization of simulation-based training system (HumRRO Final Report No. 8313). Alexandria, VA: Human Resources Research Organization.

Sticha, P.J., Schlager, M., Buede, D.M., Epstein, K., & Blacksten, H.R. (1990). Specifying skill-based training strategies and devices: A model description (ARI Technical Report 897). Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Vineberg, R., & Joyner, J.N. (1980). Instructional system development (ISD) in the armed services: Methodology and application (HumRRO Technical Report 80-1). Alexandria, VA: Human Resources Research Organization.

Zimmerman, W., Butler, R., Gray, V., & Rosenberg, L. (1984). Evaluation of the HARDMAN (hardware vs. manpower) comparability methodology (ARI Technical Report 646). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.