①

U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1653

The Optimization of Simulation-Based Training Systems: A Review of Evaluations and Validation of Rule Bases

Michael J. Singer
U.S. Army Research Institute

November 1993

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

John A. Dohme
Bruce W. Knerr

Accesion For		
NTIS CRA&I		☑
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

NOTICES

DISTRIBUTION: Primary distribution ~~of this report~~ has been made by ARI. ~~Please address correspondence concerning distribution of reports to:~~ U.S. Army Research Institute for the Behavioral and Social Sciences, ~~ATTN: PERI-.)X~~, 5001 Eisenhower Ave., Alexandria, ~~Virginia 22333-5600~~.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
1993, November

3. REPORT TYPE AND DATES COVERED
Final Jan 90 - Jun 93

4. TITLE AND SUBTITLE
The Optimization of Simulation-Based Training Systems:
A Review of Evaluations and Validation of Rule Bases

5. FUNDING NUMBERS
63007A
795
3104
H01

6. AUTHOR(S)
Singer, Michael J.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute
STRICOM, Orlando Field Unit
12350 Research Parkway
Orlando, FL 32826-3276

8. PERFORMING ORGANIZATION REPORT NUMBER
ARI Research Report
1653

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and
Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
--

11. SUPPLEMENTARY NOTES
--

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

The Optimization of Simulation-Based Training Systems (OSBATS) was designed to provide engineers involved in the concept formulation process for simulators and training devices, a tool with which to consider and tradeoff alternative features. OSBATS would enable engineers to tradeoff fidelity and instructional features in the context of a training strategy to obtain designs that minimize cost for desired performance levels. OSBATS consists of five modules: a simulator decision module, a training device media selection module, an instructional features module, a fidelity optimization module, and a resource allocation module. The heart of the OSBATS model resides in the expert system rule bases used in the instructional features and fidelity optimization modules to make decisions about instructional and fidelity features to be incorporated in the simulator designs. This report presents an overview of the OSBATS model and prototype software and summarizes efforts to validate its utility. An effort to validate the instructional features and fidelity rule bases is described in detail.

The approach used to validate the rule bases was to select example tasks and use

(Continued)

14. SUBJECT TERMS

Training system	Training system models
Optimization	Media selection
Fidelity	Instructional features

15. NUMBER OF PAGES
81

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT
Unlimited

NSN 7540-01-280-5500

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

13. ABSTRACT (Continued)

the OSBATS rules to produce recommendations for fidelity levels and instructional features. Those results were then presented to two groups of subject matter experts: instructor pilots at the Aviation School at Fort Rucker, and researchers at the U.S. Army Research Institute for the Behavioral and Social Sciences Research and Development Activity at Fort Rucker, Alabama. The two groups were queried on their agreement with the OSBATS-generated recommendations for the tasks and their rationale for assigning fidelity features and instructional features to the tasks. The results of the group interviews are summarized and discussed in the context of the previous evaluations of OSBATS. Requirements for development and fielding of OSBATS-derived software are discussed in the conclusion.

Research Report 1653

The Optimization of Simulation-Based Training Systems: A Review of Evaluations and Validation of Rule Bases

Michael J. Singer
U.S. Army Research Institute

STRICOM Orlando Field Unit, Florida
Stephen L. Goldberg, Chief

Training Systems Research Division
Jack H. Hiller, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

November 1993

Army Project Number
2Q263007A795

Training Simulation

FOREWORD

The effectiveness of a simulation-based training program is directly related to the instructional quality of the simulator, and its availability is directly related to costs of procurement, operation, and maintenance. One problem faced by the military is determining which instructional features and levels of fidelity are necessary and sufficient for the stated learning objectives. Behavioral and analytical techniques that can quickly project or predict the features required are not readily available for instructional designers. In addition, information on the effective use of training devices in courses of instruction is sparse and theories that can extrapolate from that information are weak. The development of models, databases, and techniques addressing advanced training technology will help to remedy this situation. The potential effect on the Army will be to reduce the cost of training devices and, at the same time, increase instructional effectiveness and availability.

In response to these concerns and problems, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) and the Project Manager for Training Devices (PM TRADE) [now incorporated within the U.S. Army Simulation, Training, and Instrumentation Command (STRICOM)] joined efforts [Memorandum of Understanding (MOU) for Technical Coordination, May 1983; MOU Establishing ARI Field Unit, March 1985; Memorandum of Agreement (MOA) on Advanced Technology for the Design of Training Devices, October 1991]. The ARI Task, Advanced Technology for the Design of Training Devices and Simulators, incorporated ARI's efforts under the most recent MOA and provided the framework for the work reported in this document. PM TRADE has maintained an active interest in this work and has participated in every phase of the development of OSBATS and its follow-on, the Concept Formulation Process (CFP) Aid. The OSBATS and CFP Aid models have been used by the engineering directorate at STRICOM and they have been released to a number of contractors who build training devices and simulators for the Army. These models and techniques provide the basis for supporting the integration of behavioral and engineering data, knowledge, and expertise in training device design.

EDGAR M. JOHNSON
Director

ACKNOWLEDGMENTS

The author appreciates and acknowledges the support received in conducting this work and preparing this report. John Dohme of the U.S. Army Research Institute Aviation Research and Development Activity (ARIARDA) provided research material, task data, and served as a subject matter expert. Captain Weiler of ARIARDA was most helpful in scheduling the interviews with the instructor pilots at Fort Rucker.

THE OPTIMIZATION OF SIMULATION-BASED TRAINING SYSTEMS: A REVIEW OF EVALUATIONS AND VALIDATION OF RULE BASES

EXECUTIVE SUMMARY

Requirement:

The Optimization of Simulation-Based Training Systems (OSBATS) is a prototype set of models designed to analyze training device requirements, specify training device features, and perform cost/benefit tradeoffs between design alternatives. The system consists of five modules, each of which is based on theoretical models using theoretically or experimentally based formulas, algorithms, or heuristics. Two important functions performed by OSBATS are the recommendation of instructional features and fidelity levels based on task requirements. These recommendations are made by expert system rules that are the core of each module. Verification of these two models is critical to the usefulness of this system, or any derivative system. The goal of this effort was to investigate the acceptability of the recommendations produced by the expert systems rules for instructional features and fidelity levels.

Procedure:

Initial Entry Rotary-Wing (IERW) tasks were selected as a focus for the investigation. Information about the tasks was elicited from subject matter experts at Fort Rucker, Alabama. That information was then used with the OSBATS model rules to derive a set of recommendations. The recommendations were used as the basis for two group interviews, one with instructor pilots and the other with researchers at Fort Rucker. Each interview presented the results of the OSBATS analysis and elicited discussions about the recommendations. The agreement or disagreement, with comments, is presented as a measure of the validity of the rule bases underlying the recommendations.

Findings:

The instructor pilots agreed on 70% of the task fidelity recommendations. The researchers agreed with the rule-base outcomes on 98% of the prescriptions. The instructor pilots agreed with 72% of the instructional feature recommendations,

and the researchers agreed with 77% of the assignments. This level of agreement serves as an indicator of the validity of the rules used to assign instructional features and fidelity levels to a training device configuration by the OSBATS model.

Utilization of Findings:

The results of the rule-base validation provide a basis for accepting the rule bases used in the OSBATS prototype to select instructional features and fidelity options for a small group of rotary-wing tasks. Previous analyses and evaluations investigated the credibility of data used, the modeled optimization routines, and the difficulty of gathering the internal model information used by the OSBATS prototype. The results of this and previous efforts suggest improvements for the next version of the system. The results also support the potential usefulness of the OSBATS aiding system.

THE OPTIMIZATION OF SIMULATION-BASED TRAINING SYSTEMS: A REVIEW OF EVALUATIONS AND VALIDATION OF RULE BASES

CONTENTS

LIST OF TABLES

CONTENTS (Continued)

THE OPTIMIZATION OF SIMULATION-BASED TRAINING SYSTEMS: A REVIEW OF EVALUATIONS AND VALIDATION OF RULE BASES

Introduction

Simulation-based training systems use training devices to simulate critical aspects of tasks in order to support the acquisition of specific content-domain skills and knowledge. There is considerable experience-based knowledge about how training devices and simulators can be designed, constructed, and implemented in a program of instruction or used for training. However, experimentally based information about performance and training effectiveness is limited (Hays & Singer, 1983, 1989; Sticha, Singer, Blacksten, Morrison, & Cross, 1990).

The training device specification is developed as a part of the overall acquisition process. The acquisition process structure is prescribed in Regulations and Standard Operating Procedures documents (AR 71-9, 70-1, STRICOM SOP 66). This prescribed structure details <u>what</u> has to be done in terms of meetings, decision points, and paperwork. What is not specified is the information that should be used to support design decisions, and appropriate mechanisms for conducting the required trade-offs in defining training simulations or training devices.

The need for developing organized information and models supporting the design of devices has been documented many times (Hays & Singer, 1983; Zeidner & Drucker, 1988). The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has worked for several decades to provide models for determining the cost-effectiveness of training device configurations (Zeidner & Drucker, 1988). A recent effort has produced a theoretical prescriptive model for performing trade-offs within and between features in training device concept and use. The Optimization of Simulation-Based Training Systems (OSBATS; Sticha, Blacksten, Buede, Singer, Gilligan, Mumaw, and Morrison, 1990) is a computer-based implementation of this model. The system uses task data and training program information to prescribe features and conduct cost/benefit trade-offs. OSBATS represents the first attempt to develop organized information and models that support the prescriptive design of effective training-device-based systems (Hays & Singer, 1983).

Adequate evaluation of an aiding system (like OSBATS) is designed to produce confidence in the supporting analyses, trade-offs, or recommendations made by the model(s) contained in the system. There are many possible levels of evaluation that can be performed, from formative evaluations performed during development to summative evaluations conducted after development or during implementation. There are many parallel system-oriented issues that are inherent in these evaluations: testing the system for user acceptance, usefulness, reliability, validity, effectiveness, and possible process improvement. All of these system issues have to be addressed in the context of the organization structure, the workers, and the organization's environment for change.

Given the different goals of evaluations, and the many possible problems in achieving those goals, it becomes evident that there are many ways to test or evaluate a theoretical model-based system like OSBATS. One approach (that isn't practical in the real world) would be to collect the necessary data, use the system to generate training device recommendations, build several of the recommended devices, and test for training effect while evaluating actual device development costs. A more feasible alternative is to conduct actual or theoretical tests of critical portions of the aiding system. The individual formulas, algorithms, and heuristics could be experimentally tested by building devices that vary on a single relevant dimension (e.g., motion fidelity or visual resolution) and comparing the training outcomes. Alternatively, the system could be used to "reverse validate" the design of existing systems, comparing the outcomes of the system to the actual costs and effectiveness of existing systems. More convenient, less expensive, and less rigorous evaluations can be conducted by having experts evaluate intermediate or final outcomes of the system. This is the approach that was used to investigate the rules used by OSBATS to recommend appropriate device features at specifically needed levels. In all evaluations the major question must center on what is gained by performing the tests and evaluations versus what those tests cost.

This report presents a review of the results of previous evaluations, the results of a validation effort for the rulebase component of the model, and conclusions about the validity of outcomes produced by OSBATS. The first section briefly describes the OSBATS aiding system. The second section introduces basic decision aid parameters, and then reviews previous formative and summative evaluations of OSBATS. The third section presents a comparison of the rulebase outcomes generated by OSBATS with the judgments of two groups of "experts." This is one practical method for validating complex rule bases. The section presents the method used, procedures followed, and results of a brief interview with two groups of subject matter experts (SMEs) on the adequacy of the recommendations generated by the two sets of OSBATS rules. The final section presents conclusions about the validity of the rules as implemented in OSBATS and provides recommendations for the next step in developing a system for supporting the training device concept formulation process in the U. S. Army.

OSBATS: A Decision Aid

The general need to improve information handling and decision-making in our information society has led to increased development and use of decision aids throughout all areas of government and business. Decision aids are programs that are **designed** to aid users in modeling the problem environment, supporting problem solution, or discovering potential problems (Thierauf, 1988). Decision aids have three essential parts; a **model system**, a **user interface**, and a **data base system** (Sprague & Watson, 1977). In order for a decision aid to be useful, it has to be informative, interactive, and support the user in performing a decision-making task.

OSBATS is a first step toward a decision aid for general training device design. The OSBATS program has theoretically based internal heuristics and algorithms that model the training device trade-off issues identified as important (Sticha, et al., 1990). These issues were identified during the top-down generation of a theoretical model of the training device development process. OSBATS was designed to support and aid that theoretical process and embodies that analytically developed training device design process. The prototype was developed to address the domain of rotary-wing operations training. It has a user interface that supports the selection of different modules, analyses, and screens. The OSBATS system has a model of the theoretical process. OSBATS lacks a mechanism for handling (entering or manipulating) data or the expert system rules in the different rulebases (which includes the information used by the rules to reach recommendations). The prototype software uses standard ascii data files to store the data used in the models and rules.

As a result of OSBATS being a theoretical prototype rather than specific user focused and derived software, the OSBATS models address a wide range of trade-off functions that require a wide variety of data. The model aspects of OSBATS are the strongest parts of the system because of this theoretical approach. Each of the five program modules in OSBATS (see below, and Sticha, et al., 1990) has algorithms and heuristics that perform different functions. For example, two program modules incorporate general models of learning in order to make predictions about the amount of time required to train, one at a more detailed level than the other. Also, two modules are structured around rule bases that identify required training device features. These rule bases are the focus of the evaluation effort presented later in this report.

The OSBATS user interface is a mouse-driven point and select system of menus with some keyboard input capabilities (Sticha, et al., 1990). The main structure is a menu format which presents a list of options for dealing with each module. These options can be selected by entering the highlighted letter from each option, or by placing the mouse cursor on the option and pressing a mouse button. Within each module, menu options are presented at the bottom of the screen. These options can be selected using the same techniques.

The OSBATS data base is the least developed portion of the system, consisting of a collection of simple matrix files in fixed format without a database engine or user interface (Sticha, et al., 1990). In order to enter data, the user must use a word processor to generate ASCII files that are appropriately formatted and named. The lack of database support constitutes one of the largest impediments to attempting to test and evaluate the prototype. In order to exercise the system on the target tasks, seventeen different data files have to be constructed, each with long strings of numbers and spaces representing the data available for each task. An extra space, a missing space, a missing number, or an incorrect number would invalidate the exercise, and perhaps crash the system.

The OSBATS prototype does provide a coherent methodology for addressing the trade-offs required to produce a cost-effective training device configuration (Sticha, et al., 1990). The initial formulation of OSBATS focused on rotary-wing flight training, addressing issues that center on the conditions necessary for training student pilots to fly helicopters. OSBATS consists of five modules that cluster tasks for analysis (Simulation Configuration), analyze and assign features for the tasks (Instructional Feature Selection and Fidelity Optimization), and prescribe the use of the recommended training devices (Training Device Selection and Resource Allocation).

The Simulation Configuration module clusters the task set into subsets by evaluating task characteristics, simulation needs, cost savings, and task cue - response (fidelity) requirements. The tasks are sorted into part-task and full-mission sets for further analysis. This module also labels some tasks as inappropriate for simulation-based training when safety, simulation requirements, and cost-savings are so low that training on actual equipment is more reasonable. The part-task and full mission task sets are used as the basis for analysis by the other modules.

The Instructional Feature Selection module uses rules to select instructional features that are appropriate for training tasks. The rules use task characteristics to identify Instructional Features that can improve the efficiency of training. Efficiency is achieved by reducing the time or cost required to achieve a desired task performance level on a training device. Benefit measures based on this assessment are assigned for each feature, and are weighted by the tasks addressed by each feature. That benefit value is then combined with cost estimates to form a ratio that is used to rank the instructional features value or importance for the task set. Several different sets or packages of instructional features can be prepared that address different task sets, reach different cost levels, or provide different levels of benefit. These sets are then passed on to the Fidelity Optimization module for inclusion in prospective training device configurations.

The Fidelity Optimization module uses rules to identify and rank order the appropriate fidelity dimensions and levels for task sets. The rules use each tasks' cue and response requirements to identify the appropriate fidelity dimension and level. The training benefit predicted from that level of fidelity is combined with an estimate of the associated development cost to generate a cost/benefit ratio. The cost/benefit ratio can then be used to determine the dimension levels for which the cost of increased fidelity is justified by increased training effectiveness. The output is one or more possible training-device configurations, where each has the greatest effectiveness for the given cost goal and/or different task set.

The Training Device Selection module generates a linear estimate of time that will be used in training with each device (existing and proposed), establishes a preliminary sequence of device use for each task, and conducts a simple trade-off analysis for a set of proposed devices. This analysis will show some devices to be redundant or inefficient for training the targeted tasks. The time-to-learn and cost-to-develop estimates are used to determine the device family that meets the training requirements for all tasks. A device family is a mixture of devices that can be used in sequence for training. The device family may consist of one or more low-fidelity training devices, a full-mission high fidelity simulator, and the actual equipment. The trade-off model considers the benefit and cost of the fidelity dimensions, the benefit and cost of the instructional features, the time to train, and the level of training possible on each device for each task. The output consists of a list of training devices that will train all the targeted tasks to established criteria in the minimum time.

The Resource Allocation module supports a more detailed allocation of the family of training devices selected with the aid of the Training Device Selection model. Resource Allocation differs in that it attempts to mathematically optimize the sequencing and timing of task training on each piece of equipment. The goal is to minimize the cost of achieving the specified training requirements. The difference (from the training device selection module) is that a more detailed and rigorous approach is used to account for the use of each training devices, and it allows the user to specify constraints on the time and sequence of all devices used. The output of the module is in the same form as that of the Training Device Selection module, a list of training devices that will train the tasks to required criteria. In addition, the module considers the number of students that will be trained in any year and the life-cycle of the training devices to determine the number of devices that will be needed for training.

Evaluations of OSBATS

The rationale for evaluating a decision aid is the same as for evaluating any other software product. The developer, user, and manager need to know how well the system addresses the targeted problem. There are many ways to evaluate systems, and each approach should lead toward accreditation by the intended user (Williams & Sikora, 1991). Accreditation means acceptance or certification that the system is adequate for a specific purpose. One development activity that contributes to accreditation is testing the program and making changes based on those test results as a part of the verification process. This process also involves documenting the software, performing sensitivity analyses, and verifying the information used.

The evaluations that have been conducted in the OSBATS program address the different aspects of certification introduced above. Analytical formative evaluations were performed by the developer (Sticha, et al., 1990). The formative evaluations consisted of repetitive and ongoing briefings and interviews with engineers about the models, data, and use of the system. The results were that the system; (1) addressed the concerns of multiple users, (2) was difficult to comprehend, and (3) required data that was not normally available. Ragusa, Barron, and Gibbons (1989) conducted an analytical comparison of OSBATS and another training device configuration analysis aid. Their approach was to apply OSBATS in a related domain for which it was not designed. This application identified the many problems in acquiring and organizing the required input data. An evaluation was also performed by Willis, Guha, Hunter, and Singer (1990) to determine the difficulty of gathering "internal" model information. (Internal model information is not specific to a single session with the system, but is used in many different applications [eg. feature cost or effectiveness information].) The effort was primarily concerned with the ease of expansion of OSBATS through acquiring additional information. The information was found to be available, but was difficult and tedious to acquire with any validity. Each of these efforts are briefly reviewed in the next sections.

Formative Evaluations

During the course of developing the OSBATS models, several formative evaluations were conducted (Sticha, et al., 1990). The initial efforts consisted of a structured demonstration of the model and a structured interview. These evaluations focused on the validity of the model approach, the accessibility and representation of the requirements data, the use of relevant information by the model, and the relevance or usefulness of the model outcome. After the model had reached the stage of integrated functioning, extensive formal interviews were conducted with five engineers in order to again evaluate the same aspects. Detailed results of those interviews, the conclusions, and recommendations are presented in the published report documenting the model development (Sticha, et al., 1990).

There were several major conclusions from the developers' structured evaluation of the OSBATS models (Sticha, et al., 1990). The first conclusion was that OSBATS does address many of the normal analyses that engineers are concerned with during the development process, although in different ways and with different emphasis. On this basis, it was concluded that OSBATS has reasonable face validity. The second conclusion, which weakened the first, was that OSBATS does not reflect many of the normal constraints on the development process. These include school requirements aside from the training factors addressed, the technological risk of feature development, and the time constraints on training device development. The third major conclusion was that the users were confused about the derivation of benefits and the constant use of normalization routines in the models. This seemed to indicate that the engineers did not thoroughly understand the details of the models and might disagree with some portion of the many assumptions and theoretical algorithms used in OSBATS. In addition, there was confusion about the development of a group or family of training devices, when the engineers normally developed a single device to meet all training requirements. Finally, there was considerable concern about the availability of the data used by the models. Examples of the data required and not normally available are estimates of the students prior knowledge of each task (on a 0 to 1 scale) or the amount of classroom training required for training for each task. A further limitation was the inability to inspect and alter basic cost and benefit information. The consensus was that, overall, the OSBATS prototype was not immediately usable due to differences in the level of analyses, the different scope of the analyses, and the high cost of gathering and difficulty of organizing data for different applications.

During the development of OSBATS, as a result of the formative evaluation interviews and meetings, the approach and scaling of the data used were altered to address some issues within the theoretical framework raised by prospective users. Some effort was also made to instruct the proponents (who were also the prospective users) in the broader theoretical approach. The system still could not be fielded immediately because of the lack of a data base, the lack of a large set of information for that data base, and differences between OSBATS and the customary engineering approach to training device specification.

Independent Analysis

In 1989, the Institute for Simulation and Training (IST) was tasked to investigate several cost and training effectiveness models for validity, usability, and usefulness (Ragusa, Barron, & Gibbons, 1989), one of which was OSBATS. The approach used to test OSBATS in that effort was to apply OSBATS in a domain (tank gunnery) for which it was not originally implemented (the original expert system rules were developed for rotary-wing operations training). The problems in gathering the input data and applying OSBATS were then documented. The domain required collecting and entering gunnery task characteristics and example training device descriptions into the prototype data base. The example training devices used were the VIGS (Videodisc Gunnery Simulator) and TOPGUN (arcade-like gunnery simulator using computer generated imagery).

The process of documenting the data collection, scaling, and application problems in the gunnery domain provided an overview of the general data acquisition, scaling, entry, and format problems. Analyzing the existing training devices provided insights about the difficulty encountered of defining and scaling instructional features and fidelity dimensions. For example, calculating the field of view for an eyepoint display in which the trainee looks through a single lens that replicates a gunnery sight is not a simple task. The application of OSBATS to this different domain also provided information about the flexibility of the rule systems developed for the assignment of instructional features and appropriate levels of fidelity dimensions in the rotary-wing domain. One of the problems with the analysis was that it required some "best guesses" on the part of the analysts in order to address unmet fidelity dimension questions for the tasks. Addressing those questions drove home the fact that there was not much documented technical or empirical information available about the actual cues and characteristics required for learning and performing gunnery tasks. Still, the conclusions reached for many of these tasks seemed to match the reasoning used for the gunnery related tasks addressed in the rotary-wing application (Ragusa, et al., 1989).

The major difficulty identified in applying OSBATS during this exercise turned out to be the difficulty in structuring the data files for accurate input to OSBATS. Most of the task data had to be formatted into one of thirteen ASCII files by using a word processing system. Great care had to be taken to ensure that the appropriate files were in the appropriate subdirectory and named correctly in order for OSBATS to operate on the correct data for the analysis. In addition, the rule systems required responses to each rule for each task during an online session. If the task data changed, all of the data for that task had to be re-entered by answering all of the questions again. (If the task data had not changed, OSBATS could use the results of the previous run.)

This independent application demonstrated that the collection and organization of data for applying OSBATS in an alternative domain was extremely labor intensive (Ragusa, et al., 1989), as had been surmised earlier. The exercise also indicated that the system could generalize to roughly comparable tasks in other domains (from the original rotary-wing gunnery operations tasks to tank gunnery tasks). The final conclusion of this "evaluation" of OSBATS application was that the potential for immediate use of the prototype is very limited (Ragusa, et al., 1989).

OSBATS Information Collection and Evaluation

Willis, Guha, Hunter, and Singer (1990) conducted a study to determine the effort required to collect the required internal information for the OSBATS prototype. This information is not the task or "input" data that is used during an individual analysis, but consists of the guiding "internal" information that is incorporated into OSBATS models to make assignments and recommendations about training device features. For example, information about visual resolution effectiveness, cost, or the rationale for application (generally one or more rules in the rule bases) is internal model information. The data collection effort deserves mention in connection with the evaluation of OSBATS because the effort required to support and extend such a system provides direct information about the system's potential generality and usefulness. The collection of internal model structure information documented by Willis, et al. (1990) provided an indication of the amount of work required to field a decision aid of this nature. The evaluation also provided an evaluation of the ease of extending the system to encompass other domains. In addition, the effort served to indicate where and how improvements could be made through simplifying the model, easing the data requirements, and supporting OSBATS-specific information collection.

The primary goal of the data collection effort was to select or develop a method for acquiring task and fidelity relationship information (Willis, et al., 1990). A secondary goal was to collect additional information for use in the prototype system. The first step in the project was to investigate the research literature on rotary-wing training and the documentation on existing simulators being used to train rotary-wing operations. There was insufficient detail in these sources to meet the requirements of the OSBATS models.

The next logical step was to develop the necessary data by studying the devices, their use, and the effectiveness in training. Tasks and fidelity features were selected based on a review of tasks taught and simulator fidelity or instructional features used on many of the training devices at the USAAVNC at Ft Rucker (classified devices were not included; Willis, et al., 1990). Survey instruments were generated using task lists from current rotary-wing courses and fidelity features present on selected training devices used in those courses. The surveys focused on obtaining estimates of the needed level of fidelity for training students to perform each task. The survey focused on the features of the UH-1 Cockpit Procedures Trainer (CPT) and the AH-64 Cockpit, Weapons, and Emergency Procedures Trainer (CWEPT) and the tasks that these two devices were supposed to train. These surveys were administered to the instructors using these devices to teach at the USAAVNC.

The results of this survey were analyzed to demonstrate the basis for possible additional rules for OSBATS. The analysis method was proposed as an alternative method for generating the rules used in OSBATS (or that could be used in a comparable system) to identify or select the appropriate fidelity for different aspects of the proposed training device. (The method used originally in OSBATS, and often used in constructing other rule based systems, was to have experts decompose the reasoning used in making decisions.) Basically, the authors suggested that instructor or subject matter expert agreement on structured task/fidelity relationships could be used as a basis for rules codifying the hypothesized relationship. While the rules would have immediate face validity, they would still require experimental or experiential verification.

The data collection effort (Willis, et al., 1990) identified problems in generating sufficient information for expanding and supporting OSBATS. The primary problem is that collecting the information requires a substantial investment of resources. One major aspect of that problem is that the available information about the relationship between tasks and specific feature levels is sparse. The information can be collected through structured surveys of instructors, although there is sometimes little agreement on the relationship between a task and a particular fidelity dimension.

Another of the issues addressed in validating the cost information was the level of precision required by the OSBATS trade-off models. The general rule for OSBATS is that the cost values used in the trade-off should be in the correct scale relationship, as that allows a "correct" ordering to be achieved. For example, it doesn't matter in OSBATS if the costs for two features are ten and twenty dollars or ten thousand and twenty thousand, the method maintains the relative proportionality of the costs. The highest value would be assigned a normalized value of 1.0 and the other would be assigned a value of .5. The normalized scale also maintains the relative cost relationships between features, so that the actual cost or benefit values do not matter during the cost/benefit tradeoff analyses (see Sticha, et al., 1990 for details). Of course, for final budgeting the most valid cost estimate possible is desired, because there it does matter whether the cost is ten or ten thousand dollars. A secondary problem is in collecting cost information at the individual feature level. The available historical cost information is typically at the whole device, or major subsystem level (computer, flight system, etc.) rather than at the feature level.

The major conclusion of the data collection effort was that the internally required information (e.g. feature effectiveness and costs) required to run or expand OSBATS is available (Willis, et al., 1990). The information can be acquired from school sources using the methods developed in this effort, although collection is tedious and effortful. Sufficiently detailed information can be collected and used to expand the applicability of the system. The conclusion provides support for accrediting OSBATS and derivative systems for use in identifying and specifying fidelity dimensions and instructional features.

OSBATS Rule-Base Validation

Validation is the more difficult part of the accreditation process introduced briefly above and has a host of issues that must be addressed (Williams & Sikora, 1991). In testing a prototype for validation, one common approach is to re-examine the original objectives, then determine the accuracy and adequacy of the prototype (Marcot, 1987). In Marcot's approach accuracy is the percentage of correct predictions or decisions and adequacy refers to the range of system application. Face validity is an easier test that also addresses system accuracy and adequacy, and is often achieved through developing experts consensus on the accuracy of the system (Williams & Sikora, 1991). Face validity still contributes to accreditation, even though it is not as rigorous as other approaches. The goal of the effort reported in the rest of this report was to investigate the face validity of the OSBATS rulebases through analytic interviews with qualified Instructor Pilots (IPs) at the U. S. Army Aviation Center (USAAVNC) and flight qualified researchers at the Army Research Institute Aviation Research and Development Activity (ARIARDA) at Ft. Rucker, Alabama.

The two sets of rules in OSBATS are encoded in an expert system format (EXSYS, 1985) representing the relationships between task characteristics and features required for learning to perform those tasks. These relationships were determined by several experts from the rotary-wing flight training domain and interpreted from available relevant research (Sticha, et al., 1990). The team that developed the rule sets was composed of domain experts chosen for their history of involvement in the development and use of training devices as well as their availability for extended interaction and verification of the prototype rules. Although two members of the team were experienced pilots, only one had actually taught at the U. S. Army Aviation School (Sticha, et al., 1990). The rules make the feature recommendations that are used by the OSBATS system in all subsequent analyses. Given the introductory discussion of what can be evaluated, and the prior evaluations, the next logical step in evaluating portions of OSBATS for validity was to test the feature recommendation rules.

As previously discussed, the developers could not empirically test the information structures, the rules, or the rule interactions as the rule sets were developed (Sticha, et al., 1990). Given the level of interest generated in the validity of the rules used in OSBATS, a validation effort was in order.

There are many issues involved in evaluating expert system rules (Gaschnig, Klahr, Pople, Shortliffe, & Terry, 1983), most of which reduce to evaluating how acceptably the problem is addressed. The verification process followed in developing the OSBATS rules was to test rulebase outcomes for acceptability with a group of independent experts in the same training domain. This is a common approach to rule set validation (Gaschnig, et al., 1983) which occurs iteratively and serves to establish the face validity of the current set of rules within the application domain. The key issue is that the decisions made are judged to be reliable and relevant.

One difficulty in assessing the validity of the OSBATS rulebases arises from the availability and qualification of the judging experts (a common problem). In the area of training device specification there are no acclaimed and universally recognized experts. This led to the pragmatic solution of using available instructors and researchers to judge the specified simulation requirements. At the USAAVNC, even this effort was seriously constrained by personnel workload and availability. A further complication arises from the possibility that the task analysis information used as input for the OSBATS rules differs from the expert's understanding of task characteristics. Finally, since the OSBATS rules weren't designed to be isomorphic to mental processes, even when agreement with the outcome is evident it might not indicate similarities in reasoning between the SMEs and rule bases.

Method

Tasks used in both Initial Entry Rotary-Wing training at the Aviation Center (USAAVNC) and rotary-wing experiments being conducted by ARIARDA were reviewed and a subset of eight were selected on the basis of representativeness. This domain and set of tasks was selected because OSBATS was developed with a focus on rotary-wing tasks, although neither the rules nor the cost and benefit information included in the prototype is complete for even that domain. Requirements data for these tasks was then developed and the data was analyzed using the OSBATS rules. The results of that analysis were used to generate a feature assignment list, with definitions. The list identified the fidelity and instructional features that should apply to each task, based on the OSBATS rules and task requirements data. This document formed the basis for interviews with Instructor Pilots and Researchers at Ft. Rucker that elicited comments on their agreement or disagreement with the OSBATS rules.

Tasks

The eight tasks used were all from Initial Entry Rotary Wing training (IERW Aviator Course Primary Phase Flight Training Guide). These tasks are maneuvers taught to beginning pilots as part of their initial training. The Flight Training Guide provides detailed descriptions of the maneuvers and the standards for acceptable performance. For convenience of communication, the task labels used here are Hover Taxi, Takeoff to Hover, Land from Hover, Hover Turns, Normal Approach, Traffic Pattern, Normal Takeoff, and Hovering Autorotation. The partial descriptions of the tasks that follow are intended to provide a conceptual understanding of the tasks and make the results of the study more understandable.

The **Hover Taxi** task involves learning to move on a specified heading (often with reference to a ground track) and with the aircraft oriented in the same direction. The movement is conducted at a slow speed (only a brisk walk, e.g. 3 to 6 knots) with the helicopter skids at three feet from the ground, plus or minus one foot.

The **Takeoff to Hover** task requires a smooth ascent to hover (approximately three feet above the ground) while maintaining position and orientation over the ground initiation point.

The **Land from a Hover** requires smooth descent to the desired ground point while maintaining position and orientation over the landing point.

The **Hover Turns** task requires maintaining position over a ground point, while rotating (turning the nose) in a specified direction around the vertical center of the aircraft to a specified heading.

The **Normal Takeoff** task requires a smooth accelerating ascent from a hover or from the ground into a ground referenced flight pattern while continuously adjusting to required heading, altitude, and speed.

The **Traffic Pattern** task is a continuous procedure task which requires flying a ground referenced pattern around the airport (typically rectangular). The task requires maintaining the required headings, speeds, and altitudes. This task starts with an upwind or takeoff leg and terminates with the transition into the normal approach task.

The **Normal Approach** is a continuous control task which requires bringing the helicopter down a glide path ending in a hover over a landing zone, or terminating on the ground. The task requires continuous adjustment in maintaining orientation and angle to the landing point, while continually decreasing air speed and altitude.

The **Hovering Autorotation** is a landing task procedure which requires controlling the aircraft from a three foot hover to the ground, smoothly landing after a sudden termination of power.

Subjects

Two sets of experts participated in the discussions of feature recommendations made by OSBATS for the eight tasks. One group consisted of Dr. J. Dohme (a flight qualified Research Psychologist with ARIARDA) and Mr. D. Morgan, a flight qualified Simulator Instructor Pilot. Dr. Dohme and Mr. Morgan had been using the target tasks in their research on the effect of different types of visual systems *WHAT?*. For convenience, they are referred to in the results section as researchers. The second group consisted of six Instructor Pilots (IPs) and examiners who either instructed in advanced flight (beyond the IERW course with the UH-1) and attack helicopters, conducted proficiency tests for instructor pilot qualifications, or were instrument flight examiners. They are collectively referred to in the results as IPs. (No IPs actually training the IERW curriculum were available.) These IPs all claimed to possess the necessary skills to train in the IERW course, and several had instructed in that program prior to their current assignments. However, none of the IPs were "current and qualified" in the IERW program of instruction.

Procedure

The approach used to evaluate the rules was to interview the SMEs on their agreement or disagreement with the rule outcomes. A representative task (e.g. **Hover Taxi**) would typically form the focus for the presentation of the OSBATS assigned fidelity dimension level or instructional feature. This was followed with a discussion of any tasks that had another level or feature assigned. The guidance provided to the SMEs was to discuss whether the recommendation was reasonable given the nature of the task and why there might be differences in recommendations that they would make. This cycle of introducing a fidelity dimension or instructional feature, presenting the OSBATS rule based recommendation, and discussing the SME opinion was repeated for all recommended dimensions and features. The descriptions of the fidelity dimensions and levels used in the presentations are provided in Appendix A. The descriptions of the Instructional Features presented in the results is the same as that used during the interview. The key issues discussed for each dimension or feature are presented in the results section. The SMEs were never briefed on the task information used by the rules, or the logic and structure of the rules, so that their responses would be uncontaminated by the reasoning used in OSBATS.

Results

The overall result of the discussions was that the instructors and researchers agreed moderately well with the OSBATS recommended features. The low number of subjects and the setting and time constraints for the interview session precluded any more thorough analysis than straight-forward presentation of the consensus of the participants. The following material presents the results for each fidelity dimension and instructional feature individually. First the dimension or feature will be described, as was done during the interviews. Then the information and an overview of the rationale represented by the rules will be presented (this was not done during the interviews, but is presented for the reader's comprehension). The next paragraph will provide the agreement, opinions, and rationale (if any) provided by the instructor pilots. The number of agreements, on a dimension by task basis, will be reported there. Finally, a paragraph containing the opinions and rationale of the researchers will be presented. The fidelity dimensions are presented first, followed by the instructional features. Tables 1 and 2 present the fidelity rule base recommendations for the flight tasks.

Fidelity Dimensions

Visual Resolution. Visual resolution is described by a six-point scale which is based on the distance at which a one meter square object can be discriminated. There are one hundred and forty-nine rules in OSBATS which directly or indirectly address visual resolution. Some of those rules also address the scene content and required scene texture for various task requirements, reflecting the necessary interaction among these dimensions. The key task information used by the rules centers on the need to estimate distances and altitudes (reflecting the rotary-wing origins of the rules). The rules query the user on the number of distances and altitudes to be estimated, the accuracy required for those estimates, and the objects used in making those estimates. That information is then used to calculate the size of the visual field (in arcminutes) subtended by the smallest target object at the greatest required distance. All of the

TABLE 1. RULE-BASED FIDELITY RECOMMENDATIONS FOR HOVER TASKS

	BASIC HOVER TASKS			
	TAKEOFF	TAXI	LAND	TURNS
VISUAL				
RESOLUTION	1: m^2 @ 300m	1	1	1
CONTENT	3: grnd+trees	4: cult.feat. +graphics		3
TEXTURE	1: lines & polygons		1	1
FRONT FOV	1: 40x40 deg.	1	1	1
SIDE FOV	1: 40x40	2: 40x50	2	2
POINT EFF.	NA	NA	NA	NA
AREA EFF.	NA	NA	NA	NA
PLAT. MOTION	NA	NA	NA	NA
SEAT MOTION	shaker	shaker	shaker	shaker
SOUND EFFECTS	NA	NA	NA	NA
MAP AREA	1: 5x5 km	1	1	1

NOTE: The numbers identify the level in the dimension (see Appendix A for complete descriptions).

tasks were assigned the first level in this dimension, which calls for presentation of a one meter square (m2) object at 300m distance.

The IPs agreed with the OSBATS recommendations, with the exception of the **Traffic Pattern** task. This is scored as agreement with seven out of eight recommendations. The IPs thought that the **Traffic Pattern** task might require more resolution because it requires aligning on the ground track (the path which the student pilot is supposed to follow) in order to ensure that a straight path is being flown. For the rest of the tasks, seeing a bush or large clump of grass at approximately 300m seemed adequate to them. Much of the IPs discussion on visual resolution, content, and texture revolved around the relationship between what was being presented, the detail that could be presented, and the quality of presentation. The consensus was that these three dimensions couldn't easily be thought of in isolation. This reasoning is similar to that in the OSBATS rules structure, which addresses the three dimensions with interconnecting and interacting rules.

TABLE 2. RULE-BASED FIDELITY RECOMMENDATIONS FOR FLIGHT TASKS

	NORMAL		TRAFFIC	HOVERING
	APPROACH	TAKEOFF	PATTERN	AUTOROTATE
VISUAL				
RESOLUTION	1: m^2 @ 300m		1	1
CONTENT	5: med. dense graphics+features			4
TEXTURE	1	1	1	1
FRONT FOV	1	1	3: 40x60	1
SIDE FOV	1	3: 50x60	1	1
POINT EFF.	NA	NA	NA	NA
AREA EFF.	NA	NA	NA	NA
PLAT. MOTION	NA	NA	NA	NA
SEAT MOTION	Shaker	Shaker	Shaker	Shaker
SOUND EFFECTS	NA	NA	NA	NA
MAP AREA	1	1	1	1

NOTE: The numbers identify the level in the dimension (see Appendix A for complete descriptions).

The researchers agreed with all of the OSBATS recommendations, and did not make the distinction about the need for higher resolution for the **Traffic Pattern** task. They did note that the students immediately look for size relationships in order to judge distances and establish consistent orientation, providing support for the approach used by the rules.

Visual Content. The visual content dimension addresses the background elements of the visual display such as terrain, cultural features, vegetation, buildings, and other objects. The levels of content are assigned using examples which can easily translate into required objects for presentation with adequate resolution (see above). The OSBATS assignment varied for the task set with **Takeoff to Hover** and **Hover Turns** requiring "ground plane with realistic trees" (level 3) while **Normal Approach**, **Traffic Pattern**, and **Normal Takeoff** require "medium density graphics and cultural features" (level 5). The tasks **Hover Taxi**, **Land from Hover**, and **Hover Autorotation** all called for an intermediate level of graphics and cultural features (level 4). The rules make these assignments in a fairly direct fashion. OSBATS uses the information acquired for the resolution and texture rules to identify what needs to be presented for performance of the task.

The consensus among the IPs was that the assignments seemed reasonable as they served mainly to provide cues for guiding task performance. The IPs major objection was that with low resolution the general features might not be sufficiently discriminable for use in distance estimation in the **Traffic Pattern** task. In that sense, their judgments about the kinds of features used could not be separated from the level of resolution proposed. The level of agreement was scored as seven out of eight for this dimension. (There was no discussion about teaching distance estimation in any of these tasks, in spite of the apparent importance of the skill.)

The researchers also agreed with OSBATS guidance, and they also wanted to relate the content material to resolution and texture, arguing that these factors interact. Their exception to the recommendations was that the **Land from Hover** task didn't need much in the way of features as the task was just to set the aircraft down. The key perceptual requirement is sufficient changing detail to determine the rate of closure with the ground. This was also scored as agreement with seven out of eight OSBATS recommendations. The interesting fact was that the two groups differed on their task criteria, the IPs focusing on long distance estimation requirements in a complicated task, while the researchers focused on the minimal features needed for a simpler task.

Visual Texture. Texture refers to the method used to "fill" the scene content, in order to enhance the realism of the scene content. Texture can be provided by using mathematical functions to alter the variation and detail in appearance of generated objects. A higher level of realistic texture is provided by digitizing photos of objects and using the digitized images to "fill" the generated objects. Like Visual Content, the dimension levels use descriptive examples. All of the tasks were assigned to the basic scene-construction elements (lines and polygons) level, which is the first level. OSBATS bases the recommendation of texture on the required content and necessary resolution. When great distances and complex scenes are not needed (see Visual Resolution and Visual Content, above), OSBATS presumes that no great variety in texturing is required, and therefore recommends no extra texturing.

The consensus among the IPs was that for important visual areas of interest a higher level of texture would be required, so they disagreed with all of the OSBATS recommendations as being too low. They felt that using modulating functions within basic scene-construction elements (level 2) or a small number of digitized photographs to fill basic scene construction elements (level 3) would be more appropriate. The supporting argument was that most of the tasks required judging heights or distances and therefore requires more use of texture.

The researchers agreed with all of the OSBATS generated specifications, arguing that the objects themselves are of importance, but the texture is irrelevant to learning the tasks. They disagreed with the IPs about the need for greater texturing or special "area of interest" improvements.

Front Visual Field of View. The Front Visual Field of View (FFOV) dimension refers to the area visible to the student pilot through the front cockpit display window. The levels in the dimension are identified by increasing visual fields, both horizontal and vertical (in degrees). OSBATS rules acquire and use information about the location of important content in the visual field to make the recommendation on the FFOV. The recommendation was that all of the tasks required the first level of presentation size (40 degrees vertical by 40 degrees horizontal), with the exception of **Traffic Pattern**. The recommendation for that task was the widest level assignable (40 degrees vertical by 60 degrees horizontal).

The Front and Side FOV dimensions generated the most heated and wide-ranging discussion of any issues covered. The generally held view by the IPs was that the best way to learn the visual aspects of flight was to present the entire normal field of view. Since OSBATS recommended a more restricted field of view, this was scored as complete disagreement. The IPs made many points about the need to get students to learn to look before maneuvering, and the need to build in those looking habits and skills before passing the student. The stated problem was that with smaller fields of view the student could get away with not looking, without even making the physical response of checking out the side windows, etc. The instructors were very enthusiastic about the development of helmet displays, area of focus systems, and full dome displays. They agreed that the tasks could be taught with the lesser screen displays, but noted that the student might just be learning the simulator, and that the question of transfer to actual flight was not easily predictable.

The researchers thought that evaluating the OSBATS prescriptions was difficult, for reasons similar to those expressed by the IPs. Even so, they were scored as agreeing with all of the OSBATS recommendations for this dimension. The researchers thought that the width was less important than the vertical displacement (emphasizing the downward view), and disagreed with the structure of the levels used by OSBATS on that basis. The researchers agreed with the instructors that the optimal situation for training people on these visually oriented tasks was to use the widest view possible. However, the researchers pointed out that introductory training could be accomplished using the smaller view, as recommended by OSBATS (CITATION*).

Side Visual Field of View. The Side Field Of View (SFOV) dimension refers to the area visible to the student pilot through a right side cockpit display window. The OSBATS SFOV rules begin by sharing some of the content and resolution information. They also request information from the user about the location of important content and peripheral cues for the tasks. Again, most of the tasks were assigned the same level on this dimension, a right side window of 40 degrees vertical by 50 degrees horizontal (level 2). The rules determined that **Takeoff to Hover** needed only a 40 by 40 degree right side window (level 1) and **Traffic Pattern** required a 50 by 60 degree right side window (level 3).

The IP consensus (which disagreed with all of the OSBATS recommendations) follows from the discussion about the FFOV. The IPs agreed that the side views were probably the most important feature to have in initial training because the necessity of developing the actual physical habit of checking the different views. The disagreement was that the OSBATS prescription was too small. One IP pointed out (with apparent agreement from the others) that there were several distance and clearance cues that needed to be acquired from the side view in order to learn to adequately perform the **Normal Approach** and **Traffic Pattern** tasks. The lack of adequate wrap-around or side view was seen as the most serious lack of most rotary-wing flight simulators, and was even suggested as the basis for a considerable amount of transfer inhibition.

The researchers completely disagreed with the IPs, believing that the recommendations would provide an adequate basis for training, as was the case with FFOV. However they had several comments about unconsidered aspects and relationships. They emphasized the need (partially addressed by OSBATS) for synchronization between the field of view and the motion system. (When OSBATS considers Motion, the system questions the user about the sufficiency of correlated visual cues for task performance, a tough judgment.) In addition, they were concerned that the downward visual display, coordinated and overlapped with the forward view, was insufficiently emphasized. Still, the researchers agreed that the limited fields of view recommended would be adequate for initial task training.

Point Special Effects. Point Special Effects refers to those discrete stationary or moving elements (e.g. vehicles or lights) in the scene content that can be provided by the visual system. The OSBATS rules base the recommendation on direct information about elements needed in the visual display for training students on the task. The system did not recommend using special effects in training people on any of the tasks.

The IPs agreed with OSBATS that there was no need for point special effects in training students to perform the IERW tasks. Their discussion centered on the irrelevance of discrete content special effects cues for the target tasks (see Appendix A for a description of the special effects content levels).

The researchers concurred with the recommendation. They also (in agreement with the IPs) pointed out that this fidelity dimension did not present any cues required in performing these tasks.

Area Special Effects. The Area Special Effects (ASE) fidelity dimension refers to inserting wide area elements (eg. smoke, fog, or dust from the rotor-wash) in the scene content provided by the visual system. As with the Point Special Effects rules, this recommendation is made based on input from the user about increasing the effectiveness of training through including the special effect. If a "no" answer is obtained then nothing is recommended. OSBATS did not identify a need for any area special effects for use in training the IERW target tasks to students.

The IPs went along with the OSBATS recommendations, apparently for reasons similar to those in the rules. They discussed the tasks and could not identify any aspects of the tasks that required the kind of information added to the visual display by Area Special Effects.

As with Point Special Effects, the researchers agreed with the rule base recommendation that no area wide special effects were needed. They did not add any illuminating comments on the desirability, or lack thereof, of the dimension for training.

Platform Motion. Platform Motion may be the most often discussed and least agreed upon aspect of complex flight simulators. The dimension refers to the number of degrees of freedom and magnitude of movement made by a simulator platform. The OSBATS motion rules consider several major factors and interactions in determining the need for motion in a simulator. One key aspect considered is whether any discrete motion cues are required in learning to perform the task. Another consideration is the need for acceleration cues in the relevant direction(s). Finally, consideration is made for the interaction or correlation of motion cues and the visual display. The levels in the dimension are sequenced according to expert's judgment of how increases in motion can be engineered in a motion platform. The rules in OSBATS determined that there was no requirement among the tasks that mandated any platform motion for training.

The discussion among the IPs ranged over many aspects of motion, with very little consensus on the worth of platform motion or the required amount of motion for any given task. This was interpreted as agreement with the OSBATS prescriptions when considered in conjunction with their discussion about Seat Motion (see below). Some of the instructors held that some motion was required in general flight training in order to provide the framework for transfer to helicopters. Others disagreed, emphasizing the need for just simple onset cues and coordination with the visual display. In general, about the only thing the IPs agreed on was that large scale motion would not provide any training benefit for these IERW tasks.

The researchers were in complete agreement with OSBATS rules recommendations. They pointed out that all of the necessary cues were being provided by the visual system, and that as a result platform motion wouldn't add any training value. This latter conclusion was based in research that they had conducted (*CITE*).

Seat Motion. Seat Motion consists of simulator seat force-cuing devices that operate separately from the platform motion system. There are only three levels addressed; none, seat-shaker, and G-seat. OSBATS seat motion rules draw on the needed degrees of freedom from previous motion questions. Information is also acquired about the time duration of the acceleration and its importance for training someone on the task. For the target tasks, the OSBATS rules indicated that all of the IERW tasks would benefit from the use of a seat shaker.

The IPs general discussion of motion extended into the area of seat motion before the dimension and OSBATS results could be introduced. The general agreement was that for the IERW tasks, some indication for seat of the pants acceleration was needed, as recommended by the fidelity rules. Several IPs emphasized that the acceleration onset had to be linked to the actions and visual displays. There were doubts expressed about how well a seat-shaker could provide the necessary cues. There was no consensus on the types of cues or requirements that would lead to an unequivocal assignment of motion for learning tasks. Some of the IPs introduced the concept of level of experience as a prerequisite for motion. They claimed that more experienced pilots would be thrown off in working to transition from one aircraft to another through the use of a non-motion-based simulator. It was not clear how the IPs proposed to determine the level of experience that would lead to motion-based confusion, however.

The researchers raised many of the same points discussed by the IPs, although they did not think that a seat shaker would be very important in training students on these tasks. Nevertheless, their comments were sufficient to score this as agreement with the OSBATS recommendations. One comment made was that the purpose of the seat shaker was to provide stimuli simulating the "airborne" sensation, which would serve to keep experienced aviators from getting sick (presumably from the lack of anticipated cues).

Sound Effects. The sound effects dimension addresses those sound effects associated with aircraft operation. The levels are an arbitrary ordering of the typical sounds that can be associated with rotary-wing operations. The OSBATS rules basically inquire about the auditory cues required for performance of the tasks, and whether they are correlated with other stimuli. The rules did not identify any level of sound effects as valuable for training students on the subject IERW tasks.

The IPs analyzed the tasks much as OSBATS does, checking for any sounds that would cue the initiation or guide the performance of the IERW tasks. They ended the analysis in agreement with OSBATS that there was no reason to include sound effects for these tasks.

The researchers also agreed that the tasks could be trained without any sound effects being added. However, they went on to say that in their opinion, the sound effects providing engine and skid sounds would be useful to teach basic awareness of flight power and blade pitch changes during flight. Another point raised was that hearing the skids touch down during the Land from Hover task was a good way to know that the task was completed.

Map Area. Map Area refers to the needed size of the gaming area within which the simulator visual system is capable of operating. The issue is directly addressed in the rules analysis, which asks about the amount of area needed for demonstrating and practicing each task. OSBATS identified the smallest level of area available, 5 km x 5 km, as the level to include in the recommended simulator.

The instructors agreed with the fidelity rule recommendations that very little area is needed for the IERW tasks, with the exception of **Traffic Pattern** task. This was scored as agreement with seven out of eight OSBATS prescriptions. One of the IPs raised the issue of long sight lines being required to provide adequate guiding cues for learning the **Traffic Pattern** task. The IP was not claiming that the task could not be taught in the smaller area, just that in the general course of teaching and performing the task the greater distances were very helpful for lining up the flight path. It seems reasonable to consider extending the considerations to include the need for using objects at a distance in order to guide or initiate the task, in order to identify the need for the dimension.

The researchers had no disagreements with the size of the area indicated by the fidelity rules for training on the IERW tasks. Their basic response was that a minimum area was sufficient for learning these basic tasks. They supported OSBATS with their contention that long sight distances are often not available even in the aircraft, due to haze and fog.

Instructional Features

The time available precluded in-depth discussion of all instructional features with the IPs, although all features were discussed with the researchers. However, each OSBATS recommended feature was briefly reviewed by the IPs, with discussions centering on the features value or appropriateness for training students to perform the IERW tasks. In addition, some of the features that were not recommended were also discussed with the IPs. There was no discussion about why the OSBATS rules identified the Instructional Features (IFs) as useful but discussion was elicited from the IPs about whether the feature would be useful in training people on the tasks. The presentation follows the established format; the features are first described and the OSBATS recommendation and rationale is presented, then any comments the IPs and researchers had about the feature are provided. The basic recommendations made by the OSBATS Instructional Feature rules for the IERW tasks are presented in Tables 3 and 4.

TABLE 3. RULE-BASED INSTRUCTIONAL FEATURE RECOMMENDATIONS FOR BASIC HOVER TASKS

	BASIC HOVER TASKS			
	TAKEOFF	TAXI	TURN	LAND
ADJUNCT CAI	YES	YES	YES	YES
SCENARIO CONT.	YES	YES	YES	YES
INITIAL COND.	YES	YES	YES	YES
REAL-TIME VAR. CON.	YES	YES	YES	YES
REMOTE GRAPHICS	YES	YES	YES	YES
PROCEDURES MONITOR	---	YES	---	---
TOTAL SYSTEM FREEZE	YES	YES	YES	YES
FLT. SYSTEM FREEZE	---	---	---	---
PARAMETER FREEZE	---	---	---	---
SIM. RECORD/PLAYBK	YES	YES	YES	YES
AUTO. PERF. MEAS.	YES	---	YES	YES
PERFORMANCE IND.	---	---	---	---
AUTO PERF. ALERTS	---	---	---	---
AUGMENTED FEEDBK	---	---	---	---
AUGMENTED CUES	---	---	---	---
CRASH OVERRIDE	---	---	---	---
RESET/REPOSITION	YES	YES	YES	YES
POSITIONAL FREEZE	---	---	---	---
AUTO. SIM. DEMO.	---	---	---	---
AUTO. ADAPT. TRNG.	YES	YES	YES	YES
AUTO. CUE & COACH	YES	---	YES	YES

Adjunct CAI. Adjunct Computer Assisted Instruction (CAI) provides automated instruction to students and/or instructors on the features, capabilities, and appropriate uses of the simulator. The Instructional Feature rules assign this feature when several other complex IFs have also been identified for use. The logic that the rules are based on is that when complex features (e.g. Automated Performance Measures) are used there needs to be instruction available in the training device about how to operate those features. OSBATS assigned this feature for each of the tasks in the IERW set.

TABLE 4. RULE-BASED INSTRUCTIONAL FEATURE RECOMMENDATIONS FOR FLIGHT TASKS

	NORMAL TAKEOFF	NORMAL APPROACH	TRAFFIC PATTERN	HOVER AUTOROT.
ADJUNCT CAI	YES	YES	YES	YES
SCENARIO CONT.	YES	YES	YES	YES
INITIAL COND.	YES	YES	YES	YES
REAL-TIME VAR. CON.	YES	YES	YES	YES
REMOTE GRAPHICS	YES	YES	YES	YES
PROCEDURES MONITOR	---	---	---	---
TOTAL SYSTEM FREEZE	YES	YES	YES	YES
FLT. SYSTEM FREEZE	---	---	---	---
PARAMETER FREEZE	---	---	---	---
SIM. RECORD/PLAYBK	YES	YES	YES	YES
AUTO. PERF. MEAS.	YES	YES	YES	YES
PERFORMANCE IND.	---	---	---	---
AUTO PERF. ALERTS	---	---	---	---
AUGMENTED FEEDBK	---	---	---	---
AUGMENTED CUES	---	---	---	---
CRASH OVERRIDE	---	---	---	---
RESET/REPOSITION	YES	YES	YES	YES
POSITIONAL FREEZE	---	---	---	---
AUTO. SIM. DEMO.	---	---	---	---
AUTO. ADAPT. TRNG.	YES	YES	YES	YES
AUTO. CUE & COACH	YES	YES	YES	YES

The IPs acknowledged the need for training on how to use the training device or simulator features, but were not enthusiastic about the use of CAI to deliver that training. In addition, they seemed to assume that if an instructor pilot were given the features necessary to replicate the task conditions, they could train someone to perform the task. The discussion led to scoring this feature with the IPs in complete disagreement with the rules recommendations. The conversation indicated a distinct lack of enthusiasm for any situation in which the simulator was controlling the training rather than the instructor.

In marked contrast to the IPs, the researchers thought that additional CAI on how to use the training device instructional features would be helpful in almost all cases. This was therefore scored as complete researcher agreement with the Instructional feature rules for this feature.

Scenario Control. The Scenario Control instructional feature allows the instructor to configure the simulator prior to training sessions so that specific events occur according to a pre-planned training scenario. The major considerations used by the rules in identifying Scenario Control for inclusion focus on the times when multiple tasks are being practiced or performed, tasks are difficult, stage of learning is early, and the tasks involve continuous control movements (as when controlling the cyclic in rotary-wing aircraft). OSBATS recommended the use of Scenario Control for all of the IERW tasks.

The IPs generally had no objection to this feature, agreeing that in many ways it was useful. This was scored as complete agreement with the rule base assignments. They did not have any other pertinent comments to make about the feature assignment.

The researchers pointed out that although this was a useful instructional feature for teaching emergency procedures, it was of no use in teaching the target tasks, as none of them were actually emergency tasks (Hovering Autorotation is an exercise in which the student cuts the power and rotates to the ground under control). On that basis they disagreed with the OSBATS assignment of the feature for a simulator designed for the selected IERW tasks. This was scored as complete disagreement with the OSBATS prescriptions.

Initial Conditions. The Initial Conditions feature provides the capability to preset initial environmental and vehicle dynamic parameters from a set of previously selected values with a minimum of effort. The rules in OSBATS assign this feature based on whether the task can be measured automatically, the stage of learning, and the amount of intrinsic feedback normally available, among other factors. OSBATS recommended that the feature be used for training students on all of the IERW tasks.

The IPs agreed that Initial Conditions was a useful feature for training on the IERW tasks. One IP pointed out that Initial Conditions, like many others, needed to be programmable. It seemed that he meant that the setup should be easily accomplished, and that perhaps several different setups be available for instructor selection.

The researchers agreed that being able to alter the equipment parameters and control the environment was helpful for training students to perform the IERW tasks. They pointed out that most of these features made the simulator easier to use, but seldom affected the training process for the tasks.

Real-Time Simulation Variables Control. Real-Time Simulation Variables Control (RTSVC) allows the instructor to directly insert, remove, or otherwise alter simulator variables and parameters during training exercises. Examples of variables that could be addressed include environmental conditions (e.g. wind, haze, light levels, etc.) and equipment condition (e.g. power, electricity, oil pressure, etc.). OSBATS assigns RTSVC depending on the level of the task performance standard. If the requirement is for a high level of performance then performance must be acceptable under many conditions, which implies that changing those conditions during training will make the training more efficient. OSBATS again recommended the feature for training students on all of the IERW tasks.

The consensus of the IPs was that Real-Time Control was a necessary feature for training on the IERW tasks. They maintained that the student absolutely had to be able to master the task under different environmental aspects before they could be considered to have mastered the task sufficiently.

The two researchers also agreed completely with the assignment of the instructional feature for these tasks. They regarded this feature as one of the most helpful for training general flight skills, especially because it allows in-flight changes in the environment.

Remote Graphics. Remote Graphics Replay provides the instructor with a display of ongoing student performance during the training exercise. This may be done by student station instrument replication or CRT displays of exercise status and control data (e.g. flight path over a map, deviations from flight path angles, replication of cockpit displays, etc.). This feature can be used in conjunction with the simulator record / replay feature. The OSBATS rules base the recommendation for Remote Graphics based on the requirement for situational awareness in the tasks and the students early stage of learning. OSBATS recommended this feature for training students on all of the IERW tasks.

The IPs agreed with the rule base prescription recommending the use of this feature for training on all of the target tasks. They liked Remote Graphics Replay as an instructional feature, arguing that it increased the ability of the instructor to see errors and areas of generally poor or good performance.

The researchers disagreed with the rule base prescription recommending use of this feature for all of the IERW tasks. They classed the feature as nice but not necessary for training students to perform the IERW tasks. It was not clear what they thought about the benefit that the IPs mentioned, of using the feature to provide additional feedback to the students or provide additional guidance for the instructor in training the student to perform the tasks adequately.

Procedures Monitoring. The Procedures Monitoring instructional feature provides the capability for monitoring and documenting student performance of specific procedures. The OSBATS rules assign this feature based on the need to minimize the role of the instructor and when the task has close continuous performance tolerances. OSBATS only recommended this feature for the **Hover Taxi** task.

The IPs weren't sure of the efficacy or need for Procedures Monitoring for any task. Their point was that the instructor typically hovered right over the students shoulder anyway, and so would be aware of the performance of procedures. This contrasts with the point in the OSBATS rules about minimizing the role of the instructor. In spite of this apparent disagreement about why to recommend the feature, it is scored as agreement with seven out of eight recommendations, since they wouldn't recommend the feature for the Hover Taxi task.

The researchers agreed with the IPs in that they didn't see the **Hover Taxi** task as different from the others, and didn't see the need for this feature in training on the tasks. They indicated that the feature could be useful, but just for seeing what is being done. In addition, they were not sure the feature would be used by the average instructor. This was also scored as agreement on seven out of eight OSBATS rules recommendations, again because of the disagreement over the Hover Taxi task.

Total System Freeze. Total System Freeze provides the instructor with the ability to freeze the entire exercise. It may be initiated manually by the instructor or automatically by exceeding pre-selected parameters. The rules in OSBATS tie this feature to the use of Automatic Performance measures, the difficulty of the task, need to perform task elements simultaneously, and the type of task activity (among other minor considerations). All of the tasks had this feature assigned by the analysis system.

The instructors agreed that Total System Freeze was absolutely required in order to be able to interrupt training to provide immediate corrective instruction during the lesson. An additional consideration for them was that the feature should be easily programmable for instances when an acceptable operational envelope was exceeded by the student.

The researchers also agreed with the OSBATS prescription. They thought that this feature was useful for training on these particular tasks, using the same reasoning as presented by the IPs.

Flight System Freeze. Flight System Freeze allows the instructor to stop or stabilize one or more of the flight parameters of the exercise (e.g. simplify aspects of altitude control, attitude control, acceleration, etc.). It can also be automatically initiated based on performance measures. The Instructional Feature rule set uses information about the type of activity, the stage of learning , task difficulty, and performance requirements to determine the usefulness of this feature. OSBATS did not recommend Flight System Freeze for any of the IERW tasks.

The IPs agreed with the rule base not assigning this feature for training on the IERW tasks. They just thought that it was not needed.

The researchers also agreed with the lack of assignment of this feature by the rule base. They maintained that this feature wasn't needed for training individuals on these tasks.

Parameter Freeze. Parameter Freeze provides the instructor with the capability to stabilize one or more selected parameters of the training exercise during the entire session. It may be initiated manually by the instructor or automatically by exceeding pre-selected parameters. The feature is similar to the Real-Time Simulator Variables Control feature, but operates in a simpler fashion. An example would be to cancel the wind aspect in order to simplify the simulation for a novice. The OSBATS rule set for instructional features uses the student's presumed stage of learning, the rated task difficulty, and task performance requirements to determine whether to include this feature in the simulator recommendation. The OSBATS analysis did not recommend Parameter Freeze for inclusion in the proposed training device for any tasks.

The instructors did see a need for a selective freeze feature for training on these tasks, mostly in order to control environmental effects when the student got in trouble. This marked a complete disagreement with the OSBATS prescription for this instructional feature. The discussion seemed to indicate that their preference for parameter freeze was based in the need to control environmental influences during training.

The researchers agreed with the IPs in seeing a need for a parameter freeze function in training on the target tasks. They also thought that parameter freeze might be useful, in combination with positional freeze, if there was no instructor present.

Simulator Record/Playback. Simulator Record/Playback allows the simulator or training device to record a student's actions and inputs during a training exercise. The simulator can then dynamically replay the exercise or selected segments of the exercise for the student's review. The rule system in OSBATS uses information about the type of task performance, the stage of learning, and task difficulty in assigning this feature. OSBATS recommended record/playback for use with all of the IERW task training.

The consensus among the IPs was that this feature would be useful for initial task learning on the target tasks. This was scored as complete agreement with the OSBATS prescriptions. One of the instructors thought that a better effect would be achieved from videotaping the students activities, which could be reviewed for continued instruction without the use of the simulator.

The researchers also thought that Simulator Record/Playback would be a good feature to use in training on the target tasks. Their rationale for use was that the feature could be used to illustrate student errors. This was also scored as complete agreement with the instructional feature rule outcomes.

Automated Performance Measurement. Automated Performance Measurement (APM) provides the capability to calculate quantitative measures of student performance which can be used to assess student progress and to diagnose performance problems. The Instructional Feature rules presume that if performance could be measured by the simulator that it should be. The rules also tie this feature to the stage of learning, the need for situational awareness in the task, and the presence of intrinsic feedback in equipment operation. OSBATS recommended Automated Performance Measurement for all of the tasks except **Hover Taxi**.

The IPs were skeptical about using this feature for training on any of these tasks. As a result, this was scored as agreement on only one out of eight OSBATS prescriptions. The major issue for them was whether the feature could have the capability of setting the performance standards as a student learned the tasks. The instructors felt they had the ability to progressively modify the standards and provide tailored feedback about performance to the student, which they thought a machine could not do. Other issues that they raised concerned how the measures would be used and whether they would provide any additional information to the instructor. The arguments presented by the IPs demonstrated that while they may understand rotary-wing operations, they are lacking in knowledge about training technologies and instructional strategies.

The researchers thought that Automated Performance Measurement would help instructors judge the performance of the student, and therefore would be beneficial. This was scored as agreement with the rule based prescriptions. They also indicated that if the training were primarily automated that Automated Performance Measurement would be required as part of the general approach. This logic is closely aligned with the logic used by the rule set in OSBATS, which presumes that increasing the information provided to the instructor or removing the instructor's subjective estimations of performance will lead to increased efficiency in instruction.

Performance Indicators. Performance Indicators (PI) provides flags or signals about the students' performance at specified points or steps during the exercise. The Instructional Feature rules use information about the type of activity required in task performance and whether discrete behaviors are detectable by the computer as guides for recommending this feature. OSBATS did not identify Performance Indicators as one that would aid in training students to perform the IERW tasks.

The feature and OSBATS prescription was presented to the IPs, who agreed with the prescription. There was no discussion of the feature or the prescription.

The researchers agreed with both OSBATS and the IPs on the recommendation about the use of performance indicators for training on the IERW tasks. They did not see any need for additional information on the student's performance of the tasks during training.

Automated Performance Alerts. Automated Performance Alerts (APA) is a mechanism that sets limits for performance, and produces flags or signals for the instructor when those limits are exceeded. This feature is similar to Automated Performance Measurement. An example would be signaling the instructor when the performance parameters for Hover Taxi (3 ft. plus/minus 1 ft.) were exceeded. The feature is related to Performance Indicators, but works through limits rather than sequences. The rules in OSBATS use information on the type of activities required to perform the task, the stage of learning during instruction, whether there is normally occurring intrinsic feedback during performance, and the probability of a crash (or injury) during learning. The OSBATS analysis system did not recommend this feature for any of the tasks.

The IPs agreed that the feature was not applicable for training students on the IERW tasks. This was scored as agreement with the OSBATS rules, although their reasoning differed from the apparent rationale in OSBATS. Their discussion of Automated Performance Alerts centered on the question of measuring what to them is essentially a very subjective and almost individual standard. The IPs were convinced that the simulator would not be sufficiently flexible in measuring task performance during successful learning of many task activities. An additional objection was that requiring absolute standards would slow training excessively.

The researchers were hesitant about the efficacy of Automated Performance Alerts, agreeing with OSBATS that it wasn't needed for training students on these particular tasks. They thought that it could be replaced with automatic coaching if the goal was to replace some functions of the instructor.

Augmented Feedback. Augmented Feedback provides the capability of exaggerating the normal or naturally occurring feedback to the student during practice or learning of the task. Exaggerated means that the normal feedback is increased along some dimension, for example sounds would be made louder, visuals brighter, or motions larger. Implicitly, Augmented Feedback requires a fading control for the instructor, that would allow the instructor to reduce the feedback to normal operational levels as training and testing progresses. The Instructional Feature rule set uses information about the stage of learning and the normal feedback situation to determine the need for this feature in training. OSBATS did not recommend the using Augmented Feedback for training students to perform the IERW tasks.

This feature was not discussed with the instructors. Because the feature was not discussed, it was not included in the calculations for percent agreement for the IPs.

The researchers indicated that Augmented Feedback wasn't needed for training on these tasks. In this they agreed with the OSBATS recommendations. Due to time limitations, their rationale for non-use of this feature was not elicited.

Augmented Cues. The Augmented Cues feature is similar to augmented feedback in that it provides exaggerated information, enhancing cues that are normally present for initiating or guiding actions during learning. It also requires fading control in order to reduce the exaggeration or enhancement back to normal values as training and testing progresses. The OSBATS rules for instructional feature assignment use information about task difficulty, the stage of learning, and the presence of task cues to determine the usefulness of this feature for training. OSBATS did not identify Augmented Cues for use in training students on these tasks.

The feature was not discussed with the instructor pilots. Therefore this feature was not included in the percent agreement calculations discussed below.

The researchers agreed with OSBATS as they didn't think that Augmenting Cues would be of value in training students to perform these tasks. Their rationale was not elicited due to time constraints.

Crash Override. Crash Override provides the instructor with the option of restoring the training device to the point just prior to the crash, with normal operational parameters in effect. The feature is functionally similar to but more limited than Reset/Reposition. The Instructional Feature rule set uses information about the type of task activities, and the chances of a crash (or injury) as a basis for recommending this feature. OSBATS did not select this feature as being of benefit in training students on the IERW tasks.

The IPs saw no need for using this feature on the task set, agreeing with the OSBATS prescription. This was scored as complete agreement with the OSBATS recommendations. The consensus seemed to be that it was a normal simulator feature that was minimally useful in allowing recovery after attempting more advanced flight tasks.

The researchers also saw this feature as a normal simulator feature of minimal value in training students on the entry-level tasks, agreeing with the OSBATS prescription. Their rationale was not elicited due to time constraints.

Reset/Reposition. Reset/Reposition (R/R) allows beginning or restarting the exercise from a selected input point in the exercise. The Instructional Feature rules use information about the probability of a crash or possible injury as a result of improper performance of the task. The rules also take the stage of learning and the task difficulty into account before recommending this feature. OSBATS recommended R/R for all of the tasks under consideration.

The IPs agreed that being able to reset or reposition the aircraft would facilitate training students on these tasks. This represented complete agreement with the OSBATS prescriptions.

The researchers also saw some value in having the feature in order to save large amounts of time in training. The implication was that the feature wasn't needed for training, per se, but made the instruction more efficient by enabling the on-demand repetition of relevant portions of the exercise. This was scored as agreement with the OSBATS recommendation.

Positional Freeze. Positional Freeze provides the capability of locking or stabilizing the flight parameters only, so that the aircraft maintains position or flight pattern, allowing concentration on other aspects of the task. The rule set assigns positional freeze whenever flight system freeze is identified, using the same rules and information. The rule set uses information about type or task activity, the stage of learning, the task difficulty, and task performance requirements to recommend this feature. The Instructional Features rules did not select Positional Freeze for use in training students on these tasks.

This feature was not discussed with the instructor pilots. As a result, this feature was not included in the calculated percent agreement presented below.

The researchers saw Positional Freeze as useful (in conjunction with parameter freeze) only for hover tasks during initial training. This was scored as agreement on only four of the OSBATS prescriptions since they disagreed with OSBATS lack of recommendation for the hover tasks. They caveated the objection by linking the need to more automated instructional approaches, however.

Automated Simulator Demonstration. Automated Simulator Demonstration allows programming an automated standard demonstration of the exercise, task, or maneuver that is the focus of training. The Instructional Feature rule set uses information about the type of task activities, the student's stage of learning, task difficulty, and task performance requirements to make recommendations about whether or not to use this feature. OSBATS did not recommend the use of Automated Simulator Demonstration as an adjunct in training novice pilots on these tasks.

This feature was not of interest to the instructors, in that they saw no need for it in any situation. For these tasks, it was interpreted as implicit agreement with the OSBATS outcome.

The researchers again expressed disagreement with the OSBATS prescription. They thought that the feature could be useful in demonstrating the task requirements, and would serve to replace the (standard) introductory demonstration. Therefore thy would have assigned the feature to the device prescription. This was scored as disagreement with the OSBATS rule-based recommendation.

Automated Adaptive Training. Automated Adaptive Training provides for computer-based training that adjusts to student actions and performance levels, providing more challenging levels of the task to the student. Automated Adaptive Training can be related to any of the other automated instructional features, including the augmentation of task cues or feedback. The rules in OSBATS draw on the type of activity required in the task and the need to minimize the role of the instructor in recommending this feature. OSBATS selected Automated Adaptive Training for use in training students on the IERW tasks.

The IPs did not think much of this feature, perhaps because they interpreted it as replacing the instructor. They disagreed completely with the OSBATS recommendation that the feature should be used in training students on the tasks. They raised the issue of the instructors use of subjective standards for judging task performance and making subtle adjustments in task presentation and performance measurement.

The researchers saw Automated Adaptive Training as useful in training students on the IERW tasks, agreeing with the OSBATS prescription. Interestingly, the researchers thought the feature should be used for the same reasons that the IPs thought that it shouldn't.

Automated Cuing and Coaching. Automated Cuing and Coaching supports computer-based instruction that indicates the appropriate or most salient cues for initiating and guiding performance, as well as guidance on the actions to be performed during the task. The feature can also be related to any of the other automated instructional features, including the augmentation of task cues and normal feedback. The OSBATS rules use information about the type of behavior required for task performance, task difficulty, and the stage of learning in determining whether to recommend this feature. OSBATS recommended Automated Cuing and Coaching for all of the tasks except **Hover Taxi**.

The instructors lumped this feature with the previous one, Automated Adaptive Training, claiming that the domain was too complex for automation of the training situations. This was interpreted as agreeing with OSBATS only on the negative recommendation (for Hover Taxi). One major issue raised by the IPs was that the feature would need to be easily programmable by the instructor to begin to address the needs of the student. In other words, the instructors should be able to select the cues and direct the guidance provided automatically by the feature.

The researchers noted that the feature could be beneficial in training people to perform these tasks, without commenting on the exclusion of **Hover Taxi** or the rationale for using the feature. This was interpreted as agreeing with the OSBATS recommendation on seven of the eight tasks.

Summary of Results

The IPs agreed with OSBATS Fidelity recommendations on 62 of 88 task recommendations (70%), while the researchers agreed with

TABLE 5. SME AGREEMENT WITH FIDELITY RECOMMENDATIONS FOR BASIC HOVER TASKS

	BASIC HOVER TASKS			
	TAKEOFF	TAXI	LAND	TURNS
VISUAL				
RESOLUTION	IP R	IP R	IP R	IP R
CONTENT	IP R	IP	IP R	IP R
TEXTURE	R	R	R	R
FRONT FOV	R	R	R	R
SIDE FOV	R	R	R	R
POINT EFF.	IP R	IP R	IP R	IP R
AREA EFF.	IP R	IP R	IP R	IP R
PLAT. MOTION*	IP R	IP R	IP R	IP R
SEAT MOTION*	IP R	IP R	IP R	IP R
SOUND EFFECTS	IP R	IP R	IP R	IP R
MAP AREA	IP R	IP R	IP R	IP R

IP = Instructor Pilot agreement with OSBATS rules.
R = Researcher agreement with OSBATS rules.
* > The IPs and Researchers disagreed on the reasons for motion, but agreed with the OSBATS prescriptions.

87 out of 88 OSBATS prescriptions (98%). Table Five presents an overview of the IP and researcher agreement for the basic hovering tasks, and Table Six presents the overview for the four flight tasks. The agreements indicate that nine of the eleven OSBATS prescriptions for the fidelity dimensions at least generate user-acceptable outcomes for these tasks. The higher level of agreement between the rule system and researchers (which was almost complete) is understandable given the make-up of the design group that authored the rules (see Sticha, et al, 1990). That design group consisted of flight qualified researchers with backgrounds similar to the ARIARDA researchers. The major disagreements between the IPs and OSBATS rules were over Field of View (front and side) and Visual Texture. The IPs desired the maximum visual field possible for the presentation of visual stimuli, while the researchers sided with the lesser OSBATS prescriptions. The IPs also wanted greater amounts of texture in the visual presentation than the levels recommended by OSBATS. The higher level was claimed to be necessary for acquiring sufficient cues about height or distance. Finally, the presentation of motion generated the greatest amount of discussion among the IPs. The IPs couldn't even agree with one another over the needed amount of motion, much less the

conditions driving the amount of motion to use. Their compromise consensus was to accept the OSBATS recommendation on both platform and seat motion.

TABLE 6. SME AGREEMENT WITH FIDELITY RECOMMENDATIONS FOR BASIC FLIGHT TASKS

	NORMAL				TRAFFIC		HOVER	
	APPROACH		TAKEOFF		PATTERN		AUTOROTATE	
VISUAL								
RESOLUTION	IP	R	IP	R		R	IP	R
CONTENT	IP	R	IP	R	IP	R	IP	R
TEXTURE		R		R		R		R
FRONT FOV		R		R		R		R
SIDE FOV		R		R		R		R
POINT EFF.	IP	R	IP	R	IP	R	IP	R
AREA EFF.	IP	R	IP	R	IP	R	IP	R
PLAT. MOTION	IP	R	IP	R	IP	R	IP	R
SEAT MOTION	IP	R	IP	R	IP	R	IP	R
SOUND EFFECTS	IP	R	IP	R	IP	R	IP	R
MAP AREA	IP	R	IP	R		R	IP	R

IP = Instructor Pilot agreement with OSBATS rules.
R = Researcher agreement with OSBATS rules.

The instructor pilots agreed with the OSBATS Instructional Feature rules on 104 of 144 task assignments (72%, with three features skipped due to time constraints, see Tables 7 & 8). The researchers agreed with OSBATS on 130 out of 168 assignments (77%, covering the 21 instructional features and all eight tasks). (Agreement that features should not be recommended are included in both percentages). The perceived usefulness of instructional features varied widely among the instructors. They saw a role for features that could present the tasks more efficiently, but did not see the need for the automatic-type features. The researchers emphasized that grouping automated features should occur in the context of lessening required instructor time, something the IPs found abhorrent.

TABLE 7. SME AGREEMENT WITH INSTRUCTIONAL FEATURE RECOMMENDATIONS FOR BASIC HOVER TASKS

	BASIC HOVER TASKS:			
	TAKEOFF	HOVER	TURN	LAND
ADJUNCT CAI	R	R	R	R
SCENARIO CONT.	IP	IP	IP	IP
INITIAL COND.	IP R	IP R	IP R	IP R
REAL-TIME VAR. CON.	IP R	IP R	IP R	IP R
REMOTE GRAPHICS	IP	IP	IP	IP
PROCEDURES MONITOR	IP R		IP R	IP R
TOTAL SYSTEM FREEZE	IP R	IP R	IP R	IP R
FLT. SYSTEM FREEZE	IP R	IP R	IP R	IP R
PARAMETER FREEZE				
SIM. RECORD/PLAYBK	IP R	IP R	IP R	IP R
AUTO. PERF. MEAS.	R	IP R	R	R
PERFORMANCE IND.	IP R	IP R	IP R	IP R
AUTO PERF. ALERTS	IP R	IP R	IP R	IP R
AUGMENTED FEEDBK	** R	** R	** R	** R
AUGMENTED CUES	** R	** R	** R	** R
CRASH OVERRIDE	IP R	IP R	IP R	IP R
RESET/REPOSITION	IP R	IP R	IP R	IP R
POSITIONAL FREEZE	** R	** R	** R	** R
AUTO. SIM. DEMO.	IP	IP	IP	IP
AUTO. ADAPT. TRNG.	R	R	R	R
AUTO. CUE & COACH	R		R	R

IP = Instructor Pilot agreement with rule assignments.
R = Researcher agreement with rule assignments.
** = not discussed with IPs.

Discussion

The frequency of agreement on both Fidelity and Instructional Feature assignments supports the accreditation of the rule systems in OSBATS. The focus on cues needed for initiating and guiding task learning is a user acceptable approach for structuring the fidelity specification decision. The same argument holds for the assignment of instructional features. The problems seem to center on the task characteristics used to generate those assignments. The willingness of subject matter experts to accept the OSBATS rules outcomes provides evidence that the rules capture a reasonable representation of the "truth" about the assignment of fidelity and instructional features to task training.

All of the disagreements about the fidelity prescriptions seem to be based not in differences about how to make the feature prescription, but in differences of opinion about the cue complexity needed for adequate performance of the target tasks.

TABLE 8. SME AGREEMENT WITH INSTRUCTIONAL FEATURE RECOMMENDATIONS FOR BASIC FLIGHT TASKS

	NORMAL TAKEOFF		NORMAL APPROACH		TRAFFIC PATTERN		HOVER AUTOROT.	
ADJUNCT CAI		R		R		R		R
SCENARIO CONT.	IP		IP		IP		IP	
INITIAL COND.	IP	R	IP	R	IP	R	IP	R
REAL-TIME VAR. CON.	IP	R	IP	R	IP	R	IP	R
REMOTE GRAPHICS	IP		IP		IP		IP	
PROCEDURES MONITOR	IP	R	IP	R	IP	R	IP	R
TOTAL SYSTEM FREEZE	IP	R	IP	R	IP	R	IP	R
FLT. SYSTEM FREEZE	IP	R	IP	R	IP	R	IP	R
PARAMETER FREEZE								
SIM. RECORD/PLAYBK	IP	R	IP	R	IP	R	IP	R
AUTO. PERF. MEAS.		R		R		R		R
PERFORMANCE IND.	IP	R	IP	R	IP	R	IP	R
AUTO PERF. ALERTS	IP	R	IP	R	IP	R	IP	R
AUGMENTED FEEDBK	**	R	**	R	**	R	**	R
AUGMENTED CUES	**	R	**	R	**	R	**	R
CRASH OVERRIDE	IP	R	IP	R	IP	R	IP	R
RESET/REPOSITION	IP	R	IP	R	IP	R	IP	R
POSITIONAL FREEZE	**		**		**		**	
AUTO. SIM. DEMO.	IP		IP		IP		IP	
AUTO. ADAPT. TRNG.		R		R		R		R
AUTO. CUE & COACH		R		R		R		R

IP = Instructor Pilot agreement with IF rules.
R = Researcher agreement with IF rules.
** = Not discussed with IPs.

For example, the IPs desired the maximum visual field possible for presentation of stimuli. On the same issue, the researchers agreed with the OSBATS prescription that less fidelity is required for training. The IP bias seems to come from an excessive concern for the stimuli required to **perform** the task in an operational setting rather than concern over the initial learning experience with the task. In fact, it is not clear that the IPs have ever given much thought to what is actually required to learn how to perform tasks like those used in the interview.

As might be expected, there were many digressions and excursions into related topics during the discussions about recommended fidelity for training on the eight IERW tasks. The essence of a few of these discussions is worth presentation. One major issue raised by the IPs concerned the quality of the flight model and the capability for upgrading that model in the simulator. The quality issue concerns people learning to fly the device adequately, but still learning inadequate flight skills for the actual equipment. In other words, the question is one of transfer, which is also related to the IP proposed one-to-one rule. That concept proposes that the training device or simulator should have sufficient fidelity for the trainee to learn about as much during each exercise as they would during an exercise using the actual equipment. To the IPs this means having near replication level fidelity. Upgrading means that as the vehicle changes, the training device can be easily changed. These changes might be based on mission load, weather, equipment changes, or upgrades in the aircraft. It is not clear how these issues might be addressed in the rule-based OSBATS model. It also seems that the strict application of the one-to-one rule, as proposed by the IPs, would lead to higher than necessary fidelity in the training equipment. A problem that has a long history in simulator development, and which reflects IP confusion about motion and their disagreements with the OSBATS rules prescriptions about visual fields and texture.

For the instructors, the use of instructional features depends heavily on the immediate training goals and the perceived immediate needs of the individual student in performing the task. In general, they agreed with the assignment of features that would support the instructor or ease the instructor's work load. They did not want features that could replace any part of the instructor's role. This is apparently because student-instructor interaction is considered to be the central issue in the training of pilot skills. This interaction was claimed to be especially important for rotary-wing operations training where the training is essentially one on one. They believe strongly that only an instructor pilot can judge the current proficiency of the student, arguing that a machine simply cannot rapidly measure how well the student is acquiring flight skills.

The IPs major point in this argument was that the instructor was usually providing just the essential stimuli or guidance needed by the student at that particular point in his/her learning. The point is the old one that the proficient instructor provides the feedback and guidance needed for the student to progress in optimum fashion. This argument includes the belief that the good IP can predict when the student will be able to perform adequately with very little more practice and therefore the training can be stopped. A few IPs went so far as to argue that they would pass a student (qualify the student for the task) based on their belief that the student would perform the task adequately the next time. The IPs were concerned that the automation of instructional features might remove the human instructor from the loop, that the automated features would decrease the efficiency of the training. This concern was most closely related to the integration and programmability of the instructional features.

The IPs did want programmable instructional features, mostly to put the features under the control of the instructor. The instructors were particularly concerned about automating the proficiency measurement, cuing-and-coaching, and demonstration features. Again, the central concern was that automated features would remove control from the instructor, by replacing functions normally performed by the IPs. The implication was that the instructors didn't trust the devices to train in what they thought would be the correct manner. Only one of the instructors saw a need for improving the consistency of IP evaluations and thought that this improvement could be achieved through the use of automated instructional features (e.g. automated performance measurement). For the researchers, the improvement in consistency and efficiency were major points in favor of the automated instructional features.

The researchers did not seem to share the fear of instructor de-emphasis, but seemed concerned that the use of some automated features should be linked to the use of other features, forming an instructional support package. This issue is more critical for OSBATS tradeoff routines, which are based in the assumption that the features are independent, even though the rule structure does link some of the features. The instructional feature rules, as the IPs feared, do consider (and emphasize) the replacement of instructor functions by automation. The researchers emphasized that grouping automated features should occur in the context of a demonstrated lessening (research based) of required instructor time or improvement in learning by the student.

Conclusions

There are several conclusions about OSBATS that can be drawn from the development of the prototype and the various efforts to evaluate it. The formative evaluations conducted during development supported development of a system that could perform credible tradeoffs. The independent analytical evaluations provided a basis for projecting the generality and usefulness of the prototype. The data collection and evaluation efforts established that the data required by the model is available. Finally, the comparison of rule outcomes with SME opinion provides a basis for accepting the validity of the core models in OSBATS. All that remains are the normal problems associated with implementing a computer aiding system into an organization.

Formative Evaluations

The user surveys conducted by HumRRO during development (Sticha, et al., 1990) demonstrated that, for the engineers at STRICOM, OSBATS has reasonable face validity as an analytic approach to the concept formulation process. The engineers indicated that OSBATS potentially could provide a basis for conducting reasonable and useful analyses. However, they also said that the analyses did not address the necessary issues of technological risk, development schedule, and training proponent constraints. Further, the system erroneously addressed the design of multiple training devices rather than individual training devices. The consensus of the engineers was also that OSBATS required information and data that was not normally or easily available. In addition the system was not flexible enough to handle user's differences in analyses, especially in allowing the user to inspect or change information about costs and benefits.

Independent Analysis

The IST program (Ragusa, et al., 1989) emphasized the need for programmatic task analysis guides for gathering the OSBATS specific input data. Further, the entry and use of gunnery information in an OSBATS analysis demonstrated that any use of the system requires both an integrated data base and an expert system shell for efficient transfer of information. An integrated database would support the editing of data records without the requirement of exiting the system and using an independent word processing system. A database that was integrated with an expert system shell would also relieve the problem of re-entry of rule based information. The integration would allow on-line editing of the information and automatic repetition of the rule base analyses.

OSBATS Information Collection and Evaluation

The data collection effort (Willis, et al., 1990) verified that system relevant data and information exists, but not in any organized form. This finding confirms the opinions of the engineers expressed during the formative analytical reviews (Sticha, et al., 1990). Cost information is notoriously difficult to track, categorize, and maintain. The effectiveness and application information required by OSBATS has never been adequately structured, much less collected and organized for use (Hays & Singer, 1983). There was, is, and will probably continue to be considerable difficulties associated with the availability and veracity of the complex information required by the models like OSBATS.

One problem is that in accomplishing the Concept Formulation Process (CFP) tasks STRICOM does not keep any database of the information used or generated. When training requirements are delivered, they are reviewed for understanding and completeness by different specialists (e.g. training developers, acquisition mangers, engineers, logisticians, etc.) at STRICOM. The inferences about task requirements (like visual field, level of detail and resolution, etc.) and about the desired instructional approach are not documented or questioned. One outcome is that the wide range of information used during this process is not available for the next similar requirement. The lack of historical information is a limiting factor for developing decision aids that would aid the concept formulation process (see below).

Rule-Base Validation

The major conclusion of the rule base evaluation is that the OSBATS rule systems are credible. Because the outcomes elicit general agreement, general validity in the target domain should be accepted (based on the general criteria laid out by Williams & Sikora, 1991). Based on the level of agreement evidenced by the Instructor Pilots, the assignment routines used in OSBATS for selecting features for a rotary-wing operations training device are reasonably accurate. The accuracy of the rule systems is important because those selection routines constitute the most important segment of OSBATS involved in specifying training in that domain. Whether that consensus represents an optimal assignment of features to tasks for training is not clear. What is clear is that the consensus between instructors and researchers indicates that the OSBATS rules represent the a reasonable codification of the knowledge that we currently have for making decisions about training device features.

The working assumption in this evaluation effort is that the instructors used in this evaluation are typical of Army trainers. As such, their agreement or disagreement with the outcomes generated by OSBATS provides a valid foundation for the accreditation of the rule systems. It should be noted that the IPs involved in this effort were not current and qualified for training the primary flight phase (where the target tasks are first taught). They also have biases that are revealed in some of the discussions. Army trainers are typically drawn from the pool of proficient performers of the activity (e.g. flight) and teaching is a secondary consideration. This apparently fosters the assumption that the best way to learn is to do, which provides a basis for the belief that high fidelity is required in the training simulation. The learn-by-performing belief is also reflected in the one to one (simulator exercise to flight episode) rule that was proposed during discussions. The necessity for replicating only the necessary environmental and equipment stimuli and response capability in order to acquire and then improve skills is self-evident even to the instructors. They simply disagree about the complexity and quality of the stimuli needed to learn to perform the task, because of their orientation toward a complete environment for performance.

The disagreements between the researchers and the instructors responses are unsurprising given the development process used in generating the assignment rules in OSBATS (see the introduction, and Sticha, et al., 1990.) The disagreements serve to indicate areas where there may be no consensus among the general population of expert performers, trainers, training device developers, and researchers. Among other considerations, it must be noted that the IPs were less experienced with neophyte aviators than the researchers. These differences of opinion indicate important areas for research requiring clear results that can be accepted by decision makers.

Future Development and Evaluation Issues

One major conclusion to be drawn from the various OSBATS analyses is that future and further validation of OSBATS will not get any easier. The lesson learned from this evaluation experience, and supported by available literature on the development of computer aiding systems, is that the critical test for a system be whether it is defensible, clear, and performs acceptably for the users (Stamper, 1985; Thierauf, 1988; Williams & Sikora, 1991). The literature points out that the real value of a decision aid is in providing the capability for users to more rapidly examine a wider variety of issues (Thierauf, 1988). This requires that the emphasis in developing and evaluating a decision aid be placed in terms of use as well as outcome. This has yet to be done for the OSBATS system.

The good news is that evaluating the effective use of a system is much easier than attempting to evaluate the validity of the system (Williams & Sikora, 1991). The evaluation of decision aid validity will remain a problem and continue to require analytical methods similar to those used in evaluating training devices or complicated simulation systems. Unfortunately, the level of effort required, and hence the cost, are even greater than the effort and cost of evaluating and validating complicated simulation systems.

In keeping with conclusions drawn from the information sciences literature, the development of an aiding program should only proceed when the minimal organizational criteria for success are met. One major problem in attempting to aid the concept formulation process is that the process at STRICOM (a potential user of OSBATS) is not standardized (Meliza & Lampton, 1991). This was not the problem it could have been during the OSBATS development effort because there was no attempt to capture the user-specified essence of the activity to be automated, an error that has typically created problems before (Sprague & Watson, 1977). The central criteria for successful development (Sprague & Watson, 1977; Williams & Sikora, 1991) is that the target organization supports development and that the actual users and managers are involved in the development program. The lack of historical information (for example, as at STRICOM) is also a limiting factor for developing a decision aid (Thierauf, 1988). These considerations lead to the conclusion that the attempt to aid engineers at STRICOM in the concept formulation process will probably require changing the way that some of their tasks are performed.

Final Conclusions

The final conclusions from the evaluations are that OSBATS:

has applicability to the decisions faced by an analyst in the specification of a training device configuration;

can be extended to other domains when supported by extended information collection and organization;

has models, rules and trade-off analyses that are both reliable and valid;

requires mechanisms for the efficient handling of the wide range of information used in the analyses;

will require considerable expansion of the expert system rules in order to prescribe the wide range of fidelity and instructional features.

This means that there are no real technical barriers to implementing a concept formulation process aid based on the OSBATS prototype. Only the organizationally set goals and commitment for development are required.

REFERENCES

Department of the Army. (1987). Army Regulation 71-9, Materiel Objectives and Requirements. Washington, DC.

Department of the Army. (1988). Army Regulation 70-1, Systems Acquisition Policy and Procedures. Washington, DC.

Didner, R. S. (1988, May). A Value-Added Approach to Information Systems Design. Human Factors Society Bulletin, 31(5).

EXSYS, Inc. (1985). EXSYS, Expert System Development Package. [computer program]. Albuquerque, NM.

Gaschnig, J., Klahr, P., Pople, H., Shortliffe, E., & Terry, A. (1983). Evaluation of expert systems: Issues and case studies. In: F. Hayes-Roth, D. A. Waterman, & D. B. Lenat (Eds.), Building Expert Systems (pp. 241-280). London: Addison-Wesley Publishing Co., Inc.

Hays, R. T., Jr., & Singer, M. J. (1983). Research Issues in Training Device Design: The Organization of a Data Base (ARI Technical Report 588). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A140 815)

Hays, R. T., Jr., & Singer, M. J. (1989). Simulation Fidelity in Training Device Design: Bridging the Gap Between Reality & Training. New York: Springer-Verlag.

Marcot, B. (1987, August). Testing your knowledge base. AI Expert, 2(8):42-7.

Meliza, L. L., & Lampton, D. R. (1991). Engineering Functions in Formulating Training Device Concepts (ARI Research Report 1599). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A243 227)

Micheli, G. S. (1972). Analysis of the Transfer of Training, Substitution and Fidelity of Simulation of Training Equipment (TAEG Report No. 2). Orlando, FL: Training Analysis and Evaluation Group.

Ragusa, J., Barron, A., & Gibbons, S. (1989). User's analysis and evaluation of two computer-based training system design aids: OSBATS and ASTAR (Contract No. N61339-88-G-0002). Orlando, FL: University of Central Florida Institute for Simulation and Training.

Reidel, S. L. (1988, May). User Acceptance and Field Implementation of Decision Support Systems (ARI Research Report 1477). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A200 412)

Sprague, R. H., & Watson, H. J. (1977). MIS Concepts Part I. In: W. C. House (Ed.), Interactive Decision Oriented Data Base Systems. New York: Mason/Charter Publishers, Inc.

Stamper, R. (1985). Management Epistemology: Garbage In, Garbage Out (And What About Deontology and Axicology?). In: L. B. Methie & R. H. Sprague (Eds.), Knowledge Representation for Decision Support Systems. Amsterdam, N. Holland: Elsevier Science Publishers.

Sticha, P. J., Blacksten, H. R., Buede, D., Singer, M. J., Gilligan, E. L., Mumaw, R. J., & Morrison, J. E. (1990). Optimization of Simulation-Based Training Systems: Model Description, Implementation, and Evaluation (ARI Technical Report 896). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A237 266)

Sticha, P. J., Singer, M. J., Blacksten, H. R., Morrison, J. E., & Cross, K. D. (1990). Research and Methods for Simulation Design: State of the Art (ARI Technical Report 914). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A230 076)

Thierauf, R. J. (1988). User-Oriented Decision Support Systems Accent on Problem Finding. Englewood Cliffs, NJ: Prentice Hall.

Williams, M. L., & Sikora, J. (1991, June). SIMVAL Mini-Symposium--A Report. Phalanx, 24(2).

Willis, R. P., Guha, P., Hunter, D., & Singer, M. J. (1991). The Optimization of Simulation-Based Training: Model Data Collection and Utilization (ARI Research Note 91-13). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A231 436)

Zeidner, J., & Drucker, A. J. (1988). Behavior Science in the Army, A corporate history of the Army Research Institute. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Appendix A:
Fidelity Dimension and Instructional Feature Definitions

Fidelity Dimensions. The numbers used in labeling the levels of each fidelity dimension are used in the tables presented in the text. These levels were set through the consensus of an expert panel, and are not to be taken as the last word in scaling fidelity dimensions. All that can be said about the levels are that they represent reasonable steps in increasing fidelity for each dimension.

Visual Resolution: This dimension is defined as the maximum distance on the students' visual display at which a highly discriminable object one meter square can be detected. Visual resolution is set using a 6-point scale based on these distances.

One meter square (m2) detectable at:

1) 3/10 km
2) 1/2 km
3) 1 km
4) 2 km
5) 3 km
6) 4 km

Visual Content: This refers to the background elements of the visual display such as terrain, cultural features and 3-D objects. Visual content is assigned using examples. The levels used are:

1) Ground plane with a few trees.

2) Ground plane, trees and terrain relief features.

3) Ground plane, terrain relief plus realistic configuration of trees.

4) Ground plane, terrain relief, realistic configuration of trees plus low density graphics and cultural features.

5) Ground plane, terrain relief, realistic configuration of trees, plus medium density graphic and cultural features.

6) Ground plane, terrain relief, realistic configuration of trees, plus high density graphic and cultural features.

Visual Texture: This represents the method used to "fill" the scene to enhance the realism of the scene content. Visual texture is measured using descriptive examples. The five levels are:

1) Basic scene-construction elements (lines, polygons).

2) Modulating functions within basic scene-construction elements.

3) Digitized photographs (small inventory) to fill basic scene-construction elements.

4) Digitized photographs (medium inventory) to fill basic scene-construction elements.

5) Digitized photographs (large inventory) to fill basic scene-construction elements.

Front Visual Field of View: This refers to the area visible to the student pilot through the front cockpit display window. Three levels of Front Visual Field of View are measured in terms of the displayed visual angles required.

1) 40 degrees vertical by 40 degrees horizontal,

2) 40 degrees vertical by 50 degrees horizontal,

3) 40 degrees vertical by 60 degrees horizontal.

Side Visual Field of View: This refers to the area visible to the student pilot through a side cockpit display window. There are seven levels that can be assigned.

1) Right side window of 40 degrees vertical by 40 degrees horizontal,

2) Right side window of 40 degrees vertical by 50 degrees horizontal,

3) Right side window of 50 degrees vertical by 50 degrees horizontal,

4) Right side window of 50 degrees vertical by 60 degrees horizontal,

5) Left and right side window, each 40 degrees vertical by 50 degrees horizontal,

6) Left and right side window, each 40 degrees vertical by 60 degrees horizontal,

7) Left and right side window, each 50 degrees vertical by 60 degrees horizontal.

Point Special Effects: This refers to those moving elements in the background scene content provided by the simulator's visual system. Six Point Special Effects levels were developed using examples.

1) No special effects,

2) Cultural lights,

3) Cultural lights and weapons blast,

4) Cultural lights, weapons blast, and damaged vehicles,

5) Cultural lights, weapons blast, damaged vehicles, and airborne vehicles,

6) Cultural lights, weapons blast, damaged vehicles, airborne vehicles, and moving ground vehicles.

Area Special Effects: This refers to general area elements in the background scene content provided by the simulator's visual system. Levels of Area Special Effects were developed using examples of special effects.

1) No special effects,

2) Smoke and dust,

3) Rotor wash effects.

Platform Motion: This refers to the number of degrees of movement made by the simulator platform about and along the horizontal, longitudinal and vertical axes of the simulated aircraft. Platform motion is described using a 4-point scale that places multiple degrees of movement along a continuum.

1) No platform movement,

2) Three degrees of movement,

3) Five degrees of movement,

4) Six degrees of movement.

Seat Motion: This refers to simulator force-cuing devices that operate separately from the platform motion system, including seat shaker and g-seat. Seat motion is measured using examples of seat motion.

1) No motion,

2) A seat shaker,

3) A seat shaker and a g-seat.

Sound Special Effects: This refers to those sound effects associated with aircraft operation. Sound special effects were developed using examples.

1) No audio signals,

2) Weapon firing, skid noise, and some failures,

3) Weapon firing, skid noise, some failures, and normal engine operating noise,

4) Weapon firing, skid noise, some failures, normal engine operating noise, and abnormal engine operating noises.

Map Area: This refers to the size of the gaming area within which the simulator's visual system is capable of operating. Map area was established using a seven levels that places the size of the area along a continuum.

1) 5 km x 5 km,
2) 10 km x 10 km,
3) 10 km x 20 km,
4)10 km x 30 km,
5) 20 km x 30 km,
6) 30 km x 30 km,
7) 30 km x 40 km.

Instructional Features

Adjunct Computer Aided Instruction: provides automated instruction for students on the simulator in addition to the lessons provided under instructor guidance.

Augmented Cues: provides exaggerated information (enhancing cues that are normally present) for initiating or guiding actions, for the student to use in learning. This is similar to Augmented Feedback in that it requires fading control in order to eliminate the exaggeration or enhancement back to normal values for final training and testing.

Augmented Feedback: provides exaggerated feedback to the student during practice or learning of the task. Exaggerated means that the normal feedback is increased along some dimension, for example sounds would be made louder and visuals brighter or larger. Implicitly, this feature requires a fading control for the instructor, to reduce the feedback to normal operational levels during final training and testing.

Automated Adaptive Training: provides automated training that adjusts to the student actions and performance levels, providing more challenging levels of the task for learning and improved performance. This can be related to any of the other automated instructional features, including the augmenting of task cues and feedback.

Automated Cuing and Coaching: provides computer based instruction that indicates the proper cues for initiating and guiding performance, as well as guidance on the actions to be performed during the task. This feature can also be related to any of the other automated instructional features, including the augmenting of task cues and normal feedback.

Automated Performance Measurement: is the simulator capability to calculate quantitative measures of student performance which can be used to assess student progress and to diagnose performance problems.

Automated Performance Alerts: provides the instructor with a mechanism that sets limits for performance, and produces flags or signals for the instructor when those limits are exceeded. This is related to Performance Indicators, but works through limits rather than sequences.

Automated Simulator Demonstration: provides an automated demonstration of the exercise, task, or maneuver to be learned by the student.

Crash Override: provides the instructor with the option of restoring the system at the point of the crash, with normal operational parameters in effect. The feature is functionally similar to but more limited than Reset/Reposition.

Initial Conditions: provide the instructor the capability to preset initial environmental and vehicle dynamic parameters from a set of previously established values with a minimum of effort.

Flight System Freeze: provides the instructor with the capability of stabilizing or holding one or more of the flight parameters of the exercise constant. It can also be automatically initiated.

Parameter Freeze: provides the instructor the capability to stabilize a limited number of selected parameters of the training exercise for the purpose of training. It may be initiated manually by the instructor or automatically by exceeding pre-selected parameters.

Performance Indicators: provides flags or signals about the students performance at specified points or steps during the exercise.

Procedures Monitoring: provides the instructor the capability to monitor and document performance of specific procedures from a display.

Positional Freeze: provides the capability of freezing or stabilizing the flight parameters only, so that the aircraft maintains position, allowing concentration on other aspects of the task.

Real-Time Simulation Variables Control: provides the instructor the capability to insert, remove, or otherwise alter simulator variables and parameters during training exercises.

Remote Graphics: This provides the instructor with a display of current student performance during the training exercise via student station instrument replication and/or CRT displays or exercise status and control data.

Reset/Reposition: provides the capability to begin or restart the exercise from a selected or input point in the exercise.

Scenario Control: provides the instructor with capability to configure and to control the simulator so that simulated events occur according to a pre-planned specific training scenario.

Simulator Record/Playback: is the simulator capability to record the simulator conditions, student's actions and responses, and the instructor interventions during a training exercise. This allows the simulator to dynamically replay the exercise or selected segments of the exercise for the instructors or students review.

Total System Freeze: provides either the instructor or an automatic process (e.g. Adjunct CAI or Automated Performance Measurement) with the capability to freeze the entire exercise. It may be initiated manually by the instructor of automatically by exceeding pre-selected parameters.

Appendix B:
OSBATS Task Results Survey

Task 1001 - HOVER TAXI

The OSBATS fidelity assignments & descriptions are as follows:

Visual_Resolution	m2 at .3km	Distance at which a standard-sized unit can be perceived
Visual_Content	Generic features	Density of the visual scene content
Visual_Texture	Lines & Polygons	Degree of texturing of visual scene objects
Visual_Front	40x40 degrees	Visual angle in horiz. and vert. dimensions of front FOV
Visual_Side	40x50 degrees	Visual angle in horiz. and vert. dimensions of side FOV
Point_Effects	none	Level of special effects at a point in visual display
Area_Effects	none	Level of special effects across areas of visual display
Platform_Motion	none	Number of degrees of freedom in platform motion
Seat_Motion	Seat Shaker	Degree of force cuing on training device seat

Sound_Effects Complexity of sound effects available	none
Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y

OSBATS Task Results SURVEY

Task 1002 - Takeoff to Hover

OSBATS fidelity assignments & descriptions:

Visual_Resol. Distance at which a standard-sized unit can be perceived	m2 at .3km
Visual_Content Density of the visual scene content	Plane w/ trees
Visual_Texture Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front Visual angle in horiz. and vert. dimensions of front FOV	40x40 degrees
Visual_Side Visual angle in horiz. and vert. dimensions of side FOV	40x40 degrees
Point_Effects Level of special effects at a point in visual display	none
Area_Effects Level of special effects across areas of visual display	none
Platform_Motion Number of degrees of freedom in platform motion	none
Seat_Motion Degree of force cuing on training device seat	Seat shaker

Sound_Effects Complexity of sound effects available	none
Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	
Performance Indicators	
Procedure Monitoring	Y
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	

OSBATS Task Results SURVEY

Task 1003 - Land from Hover

OSBATS fidelity assignments & descriptions:

Visual_Resol.	Distance at which a standard-sized unit can be perceived	m2 at .3km
Visual_Content	Density of the visual scene content	Generic Features
Visual_Texture	Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front	Visual angle in horiz. and vert. dimensions of front FOV	40x40 degrees
Visual_Side	Visual angle in horiz. and vert. dimensions of side FOV	40x50 degrees
Point_Effects	Level of special effects at a point in visual display	none
Area_Effects	Level of special effects across areas of visual display	none
Platform_Motion	Number of degrees of freedom in platform motion	none
Seat_Motion	Degree of force cuing on training device seat	Seat shaker

Sound_Effects Complexity of sound effects available	none
Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y

OSBATS Task Results SURVEY

Task 1004 - Normal Approach

OSBATS fidelity assignments & descriptions:

Visual_Resolution Distance at which a standard-sized unit can be perceived	m2 at .3km
Visual_Content Density of the visual scene content	Medium Density
Visual_Texture Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front Visual angle in horiz. and vert. dimensions of front FOV	40x40 degrees
Visual_Side Visual angle in horiz. and vert. dimensions of side FOV	40x50 degrees
Point_Effects Level of special effects at a point in visual display	none
Area_Effects Level of special effects across areas of visual display	none
Platform_Motion Number of degrees of freedom in platform motion	none
Seat_Motion Degree of force cuing on training device seat	Seat shaker

Sound_Effects none
Complexity of sound effects available

Map_Size 5x5 km
Size of the simulations's terrain data base

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y

OSBATS Task Results SURVEY

Task 1005 - Traffic Pattern

OSBATS fidelity assignments & descriptions:

Visual_Resolution	Distance at which a standard-sized unit can be perceived	m2 at .3km
Visual_Content	Density of the visual scene content	Medium Density
Visual_Texture	Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front	Visual angle in horiz. and vert. dimensions of front FOV	40x60 degrees
Visual_Side	Visual angle in horiz. and vert. dimensions of side FOV	50x60 degrees
Point_Effects	Level of special effects at a point in visual display	none
Area_Effects	Level of special effects across areas of visual display	none
Platform_Motion	Number of degrees of freedom in platform motion	none
Seat_Motion	Degree of force cuing on training device seat	Seat shaker

Sound_Effects Complexity of sound effects available	none
Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y

OSBATS Task Results SURVEY

Task 1006 - Normal Takeoff

OSBATS fidelity assignments & descriptions:

Visual_Resolution Distance at which a standard-sized unit can be perceived	m2 at .3km
Visual_Content Density of the visual scene content	Medium Density
Visual_Texture Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front Visual angle in horiz. and vert. dimensions of front FOV	40x40 degrees
Visual_Side Visual angle in horiz. and vert. dimensions of side FOV	40x50 degrees
Point_Effects Level of special effects at a point in visual display	none
Area_Effects Level of special effects across areas of visual display	none
Platform_Motion Number of degrees of freedom in platform motion	none
Seat_Motion Degree of force cuing on training device seat	Seat shaker

Sound_Effects Complexity of sound effects available	none
Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y

OSBATS Task Results SURVEY

Task 1007 - Hover Turns

OSBATS fidelity assignments & descriptions:

Visual_Resolution Distance at which a standard-sized unit can be perceived	m2 at .3km
Visual_Content Density of the visual scene content	Plane w/ trees
Visual_Texture Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front Visual angle in horiz. and vert. dimensions of front FOV	40x40 degrees
Visual_Side Visual angle in horiz. and vert. dimensions of side FOV	40x50 degrees
Point_Effects Level of special effects at a point in visual display	none
Area_Effects Level of special effects across areas of visual display	none
Platform_Motion Number of degrees of freedom in platform motion	none
Seat_Motion Degree of force cuing on training device seat	Seat shaker

Sound_Effects Complexity of sound effects available	none

Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y

OSBATS Task Results SURVEY

Task 1008 - Hover Autorotation

OSBATS fidelity assignments & descriptions:

Visual_Resolution Distance at which a standard-sized unit can be perceived	m2 at .3 km
Visual_Content Density of the visual scene content	Generic Features
Visual_Texture Degree of texturing of visual scene objects	Lines & Polygons
Visual_Front Visual angle in horiz. and vert. dimensions of front FOV	40x40 degrees
Visual_Side Visual angle in horiz. and vert. dimensions of side FOV	40x40 degrees
Point_Effects Level of special effects at a point in visual display	none
Area_Effects Level of special effects across areas of visual display	none
Platform_Motion Number of degrees of freedom in platform motion	none
Seat_Motion Degree of force cuing on training device seat	Seat shaker

Sound_Effects Complexity of sound effects available	none

Map_Size Size of the simulations's terrain data base	5x5 km

Instructional Feature assignments & descriptions

Automated Performance Measurement	Y
Performance Indicators	
Procedure Monitoring	
Automated Performance Alerts	
Augmented Feedback	
Augmented Cues	
Record/Playback	Y
Total System Freeze	Y
Remote Graphics Replay	Y
Initial Conditions	Y
Scenario Control	Y
Crash Override	
Reset/Reposition	Y
Parameter Freeze	
Flight System Freeze	
Positional Freeze	
Real-Time Simulation Variables Control	Y
Automated Simulator Demonstration	
Adjunct CAI	Y
Automated Adaptive Training	Y
Automated Cuing and Coaching	Y