Research Problem Review 78-26

AD A076658

PSYCHOLOGICAL FIDELITY IN SIMULATED WORK ENVIRONMENTS

LEVEL II

Proceedings of a Symposium at the
Annual Meeting of the
American Psychological Association,
August 1978, Toronto, Canada

UNIT TRAINING AND EVALUATION SYSTEMS TECHNICAL AREA

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

December 1978

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

JOSEPH ZEIDNER
Technical Director

WILLIAM L. HAUSER
Colonel, US Army
Commander

Army Project Number
16 2Q162722A764

Army Training Technology

14

ARI- Res[redacted] Problem Rev[redacted]-78-26

6

PSYCHOLOGICAL FIDELITY IN SIMULATED WORK ENVIRONMENTS

Edited by Donald E. Erwin

10

Proceedings of a Symposium at the
Annual Meeting of the
American Psychological Association,
August 1978, Toronto, Canada

Frank J. Harris, Chief
UNIT TRAINING AND EVALUATION SYSTEMS TECHNICAL AREA

Accession For	
NTIS GRA&I	☑
DDC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

11 Dec[redacted] 1978

12 62

Research Problem Reviews are special reports to military management. They are usually prepared to meet requests for research results bearing on specific management problems. A limited distribution is made--primarily to the operating agencies directly involved.

408 010 A LB

TABLE OF CONTENTS

B

1.0 INTRODUCTION

Frank J. Harris
U.S. Army Research Institute

To conduct all military training using the entire system that an individual will ultimately be involved with in his operational mission is logistically impossible, cost-wise prohibitive, and probably not the most efficient way to train anyway.

Designers of training systems are faced with the problem of determining how much of the physical aspects of a system must be faithfully reproduced in order to elicit sufficient psychological realism that the training is meaningful and will indeed transfer to the ultimate operational mission.

Put another way, given that it is impossible and probably even undesirable to reproduce the complete operational environment in a training setting, how does the training designer find that minimal or optional level of physical fidelity to achieve an acceptable level of psychological fidelity sufficient to elicit the behaviors or performance levels specified as training goals or objectives?

The question of course arises as to how the training system designer decides when psychological fidelity has been achieved. What sort of analyses of psychological factors are required so that the training system designer knows when he or she has enough physical fidelity to elicit the behaviors critical for training? And when these behaviors are elicited, how does the training system designer know that there is indeed psychological equivalence between the simulated and the operational environment?

We have asked paper presenters to describe how psychological fidelity requirements and/or behavioral contingencies were isolated in the operational environment for simulation in the training environment; what training goals or performance objectives were established prior to or during simulation design; how performance evaluation or assessment was handled in the context of simulation; and what sort of cost/benefit tradeoffs seem to exist in focusing on achieving psychological fidelity rather than full-scale simulation for the work environment with which they have been involved.

Our paper presenters come from the Air Force, industry, and the Army, and their simulation interests range from individual to collective training in both machine-ascendant and what we might call "man-ascendant" systems, in which system performance is a function of man-man as well as man-machine interactions.

2.0 THE CONCEPT OF PERFORMANCE EQUIVALENCE IN TRAINING SYSTEMS

W. G. Matheny
Life Sciences, Inc.

Our topic of concern is psychological fidelity in simulated work environments. Most often when we concern ourselves with psychological fidelity we are working in the area of training. More specifically, we are generally concerned that a training device bear some fidelity relationship to some real target system. However, we may be concerned with the psychological fidelity between two situations for purposes of selection and classification; system design or in studying the effect of such operator variables as age, sex, fatigue, stress, level of experience, or other biographical, physiological or psychological descriptors using a simulated system.

However, usually we use the term psychological fidelity in the context of training and our concern is with bringing about a psychological realism rather than striving for physical realism between the two systems. I will be speaking primarily about fidelity for training in this presentation, although the fact that we may be interested in psychological fidelity for a number of other purposes should be borne in mind.

When we address ourselves to the question of the psychological fidelity of two systems, the implication is that we are looking for some way in which the two systems are alike and that this "alikeness" is some characteristic of a human within the system. The term "fidelity" as meaning the accuracy of a description, a translation or a reproduction (a la Webster) is quite unambiguous and useful. The term "psychological" is the one which gives us the problem and which must be clarified and defined.

By the term psychological we mean some behavior on the part of the operator of the system which we may observe and measure and relate to something meaningful. As a first step toward a definition, let us say that there is some psychological fidelity relationship holding between two systems when the two elicit the same behavior on the part of the human operator. Let us then talk about behavioral fidelity rather than psychological so that we are steered toward looking for and operationally defining some observable behavior. There is then behavioral fidelity between the two systems or situations when they both, under the same circumstances, elicit the same operator behavior.

Since we are trying to move toward designing systems in terms of their behavioral rather than physical fidelity, we may take our cue from psychophysics and seek to define the parameters of interest on a behavioral continuum rather than the physical. However, practically speaking, for most training systems, we must stay rooted in the physically measurable in order that we may design and build, although our concern is with the behavioral concomitants of these physical parameters. This rooting in the physical world is necessary in order to translate behavioral fidelity into physical reality. The important point is that the behavioral requirements dictate the physical characteristics rather than some perceived physical fidelity dictating behavior. Parenthetically, the definition of the physical continuum underlying the behavioral is not always easy since they must be stated in terms meaningful to the design engineer and allow him flexibility in exercising his ingenuity in bringing into being hardware which will satisfy the behavioral requirements. This definition of the physical stimulus in terms meaningful to the engineer and which allows him to exercise his art in bringing the stimulus into functional hardware is the key to the symbiotic relationship between psychologist and engineer spoken of by Col. Dan Fulgham in his recent paper before the Royal Aeronautical Society Flight Simulation Group (7).

We have said that psychological fidelity implies some sort of behavior exhibited as some measurable performance. We may then define the performance we are interested in as being the metric for measuring degree of fidelity. Again taking our cue from psychophysics, while we stay rooted in the physical measurement of certain parameters, our continua of interest are those behavioral concomitants of variations in the physical since it is human operator behavior which we want to understand and to be able to predict. From this premise I want to examine psychological fidelity in the context of a simulator of an aircraft. Certain generalizations might be drawn from this specific case which may be reasonable and useful.

Taking our cue again from psychophysics, the problem that faces us is one of systematically varying the physical environment and noting its effect upon some behavior. Unfortunately, an aircraft presents a tremendously complex set of stimuli to the operator so we are faced with the question of what realities to attend to and what stimuli to vary. We have in fact turned, in the past, to psychophysical studies in seeking an answer to the psychological fidelity of simulation and have researched such questions as "what is the threshold for perceived motion in either vestibular or the total body movement sense." We have tended to reject these findings for the most part on the basis that the isolation and variation of single stimuli within the total complex of what is going on in an aircraft is unrealistic and our object representation is not valid.

Our problem then is to identify some behavior which is reflective of the total complex of the stimuli in the aircraft control situation which is measurable and meaningful. We then must determine what we can point to in the real physical world which made the behavior happen.

One of the persistent problems in simulator fidelity is that of reproducing the requirements for closed-loop control of the system. Closed-loop control or tracking is a major part of the task of the operator of most man/machine systems, particularly aircraft, and the problem of getting the simulator to "fly like" or "handle like" the target system is a major concern. Let us take this closed-loop control aspect of the task as one for which we would like to establish a measure of behavior and from there work toward how we might define behavioral, ergo, psychological fidelity.

Let us make clear that it is operator behavior we are interested in as opposed to total system behavior. In an aircraft, for example, we are interested in some measure of the operator's output rather than such measures of total system output as attitude, airspeed, and the like. Our interest in the operator's behavior centers on his input to the controls since this is the medium by which he communicates commands to the machine. Without attempting to identify at this time the whole myriad of stimuli presented to the operator, let us assume that the aircraft system is so designed that, if we wish it to act in a certain way, particular inputs must be made into the controls. Any controller, mechanical or human, in order to produce a particular machine system output must make control inputs of a given nature. The human operator "talks" to the machine through the controls and we may focus upon this as a means of measuring the behavioral requirements of the machine behavior. We may even vary certain physical dimensions or characteristics of the system to determine whether this behavior changes and then establish a psychophysical relationship, if you will, between physical and behavioral continua in the basic psychophysical sense.

The central question is how do we measure this operator behavior. In the aircraft industry we have been measuring it for years through pilot opinion, that is, a verbal expression of the goodness or badness of the handling qualities of the system. With flight simulators the operator has been asked to compare his behavior in controlling the aircraft to the behavior he exhibits in controlling the simulator and then to express his opinion as to the degree to which they are alike. In order for the measurement to be as reliable and discriminating as possible, we have selected the most experienced operator of the system possible and supplied him with rating scales to assist him in objectifying his opinions. This behavioral measure has not been entirely satisfactory. However, it does look at the system from the behavioral point of view.

Another method by which we have asked whether two systems are alike in behavioral terms is to train individuals in a system and then transfer them to a target system and to measure how long it takes him to adapt to the new system. This method seems to me to be really answering a different question. In it we are asking what skills and knowledges gained prior to the observation of the behavior are utilized in the target system. More importantly, we are asking how long does it take the operator to learn to adapt these skills and knowledges to the target system. The measures we take in such a situation have traditionally been those of the total system output such as the aircraft altitude, airspeed, and so forth that we mentioned earlier. It does not seem to me, however, that this is a particularly suitable method for establishing the behavioral fidelity of one system versus another. It is indeed the only way we have of establishing the effectiveness of a given system for training in the target system for the particular level of experience of the individuals in which we are interested. Generally, in transfer of training experiments we are interested in the effectiveness of the total complex of the training tool and methods for reducing training time in the target system.

The fidelity question is only one part of the total question being asked with respect to the effectiveness of a training system. In asking the fidelity question we are in essence asking, is there some way in which we can come to some determination of the faithfulness with which the simulated situation represents the target situation. In this we seem all to be tacitly assuming that if we can establish a behavioral equivalence between our training situation and the target situation that we can substitute training in the simulated situation for that in the target situation with some degree of confidence. We somehow believe that if a simulated situation or system requires the same behavior as the target system, then training in that system should be beneficial and effective. At the same time we somehow believe that the physical realism and fidelity which we are now building into many systems goes beyond what is necessary to bring about behavioral fidelity.

We are continually faced with the very real question in building training systems that we must design and build some hardware components of that system which provide practice for the student. Whether the training system within which this hardware is embedded is efficient and effective or not depends upon many other variables apart from the fidelity of the training hardware. Thus, a design engineer, for example, is often asked to build such a piece of hardware with very little guidance as to its training requirements other than it is for a particular target system. On the other hand, he may receive very explicit and detailed guidance and specification with respect to the certain aspects of the hardware involved. Generally, he has little guidance with respect to the way in which he must design in order to create particular specified behaviors on the part of the operator. When he is asked to design in accordance with Military Standard 1472, he has some assurance that his anthropometric and knobs and dials design will result in certain behavior. However, when he tries to design an aircraft simulator so that it will produce the appropriate control behavior, he has practically no knowledge of the limits to place on his physical design. This is particularly true for the area of aerodynamic equations, visual scene, and motion characteristics.

How might we attack this whole problem, then, of providing information to a designer such that he could use his ingenuity to provide effectively the necessary and sufficient hardware to bring about the desired behavior, i.e., the behavior that would be exhibited in the target system? First, we must define the behavior and how we might measure it.

As shown in Figure 2-1, we can take as the behavior of interest that of the output of the operator into the controls of the system since this is a manner in which he communicates to the system and which we will assume has a characteristic signature in order to bring about particular outputs of the system. The control engineer, when designing an electro-mechanical control system, seeks to "match" the system's dynamic characteristics in order to model a control device. Methods for describing the way in which the human controller "matches" the system's dynamics and exercises control are not fully determined. The present state of this area of investigation has been summarized very professionally by Pew (14). We will come to a more specific operational definition of the operator behavior of interest to us after a bit more background discussion.

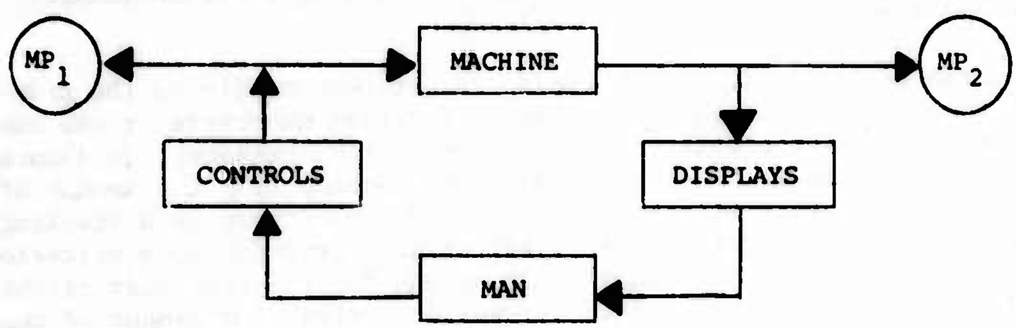


Figure 2-1. Performance measurement points (MP) in the man/machine system.

Craik in 1946 (3); Benepe, Narasimhan, and Ellson in 1954 (1); and Fitts, Bennett, and Bahrick in 1956 (5) summarized the control input signature of the operator through power density spectrum analysis and autocorrelational techniques. While Fitts was interested in understanding the nature of skilled motor behavior, Benepe et al. were more interested in a description of the operator for system design purposes. Suggestions from these studies formed the basis for my early interest in this area. For example, Fitts indicated that the cross-correlation function provided a measure of the average lead or lag of the response relative to the stimulus in the tracking situation, that is, the measure was sensitive to physically measurable changes in the system. Benepe et al. showed that while the power density spectral composition varied with practice, after attaining skill in the system the operator exhibited stationary control behavior. This conclusion appears at our present state of knowledge to be an oversimplification. However, in Benepe's experiment it is important to note that stationary behavior occurred in that situation in which the visual signal was accompanied by a synchronized auditory signal. He suggested that we cannot consider the human operator as an isolated, single channel component of a system since, when the operator was obtaining information from more than one channel, he was able to maintain highly consistent behavior as indicated by the distribution of power in his control output.

My own interest in the fidelity of simulation problem originated as a result of complaints I received from pilot trainees during World War II directed at the old "blue box" Link instrument trainers. This interest came to focus in 1950 when Williams and I carried out a survey in which we asked the question "what is the necessary and sufficient hardware to be built into a trainer in order to provide effective training for a particular system." The answer at that time was "we don't know and it costs but little more to add the hardware for any characteristic, so why worry." That was our answer. The same question has been raised over the last 27 years with essentially the same answer, and we are addressing it here again today.

In the early fifties we attempted to get some handle on the problem in both the laboratory and in very simplified maneuvers in the simulator using transfer of training as the evaluative measure. In laboratory experiments carried out by Muckler and Matheny (13) the amount of friction and direction of movement of the control forces in a tracking task were varied to determine their effect upon transfer to a criterion task. It was found for original learning and for transfer that neither differing amounts of control friction nor direction of movement of the control had a differential effect upon transfer when these variables were studied as independent variables. The interesting effect, however, came about when both control force and direction of movement were changed and negative transfer of training occurred. This has significant implications for the study of behavioral fidelity using minute and isolated aspects of the task.

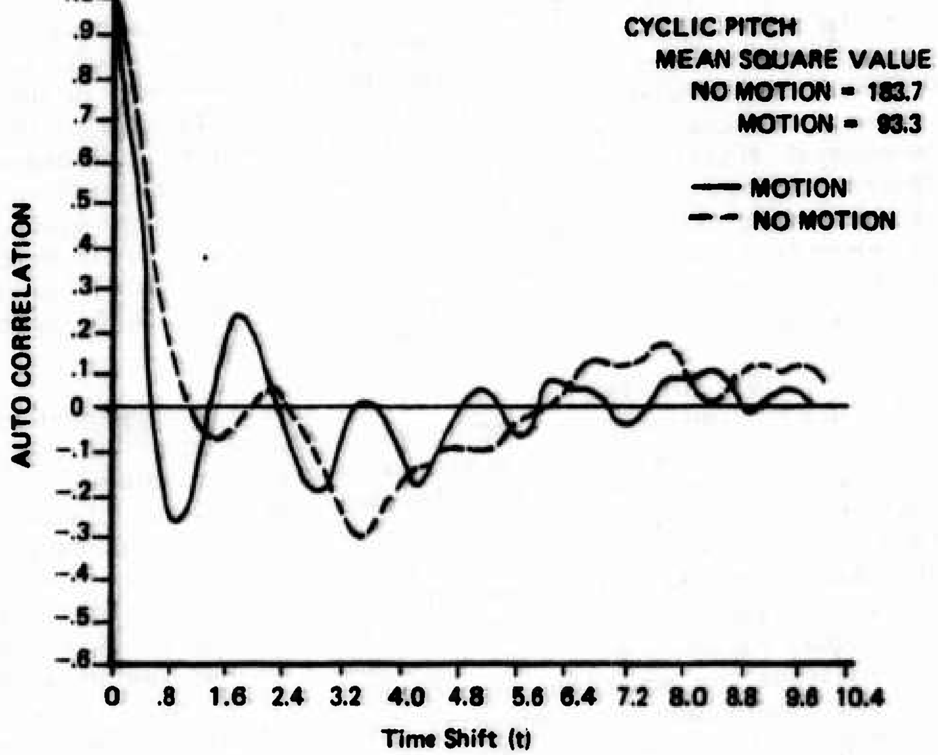
We later studied the effect of varying control forces on the transfer of training between the simulator and the AT-6 aircraft. The findings showed that both for time to reach asymptote in original learning and for transfer of training no differential effects of control force were found for the control forces studied.

Later in 1957 and 1958 in the conduct of the Army/Navy instrumentation program, the Army and the Navy supported the construction of a simulator at Bell Helicopter for use in instrumentation studies. We were not only interested in whether or not different designs brought about different control behaviors on the part of the operator but were also interested in isolating insofar as possible the cues used by the operator in carrying out his task. We were, therefore, able to obtain data on operator inputs performing the hovering maneuver in the simulator both with and without motion. The autocorrelations of the outputs shown in Figure 2-2 represent in a simplified way the essence of the behavioral fidelity approach for determining the requirements for simulators. Figure 2-2 illustrates the shift in operator control behavior under conditions of motion and no motion which could be considered a behavioral index of change in the physical system. Under other conditions of system dynamics, external forcing function or system limits no such shift in control behavior might be found.

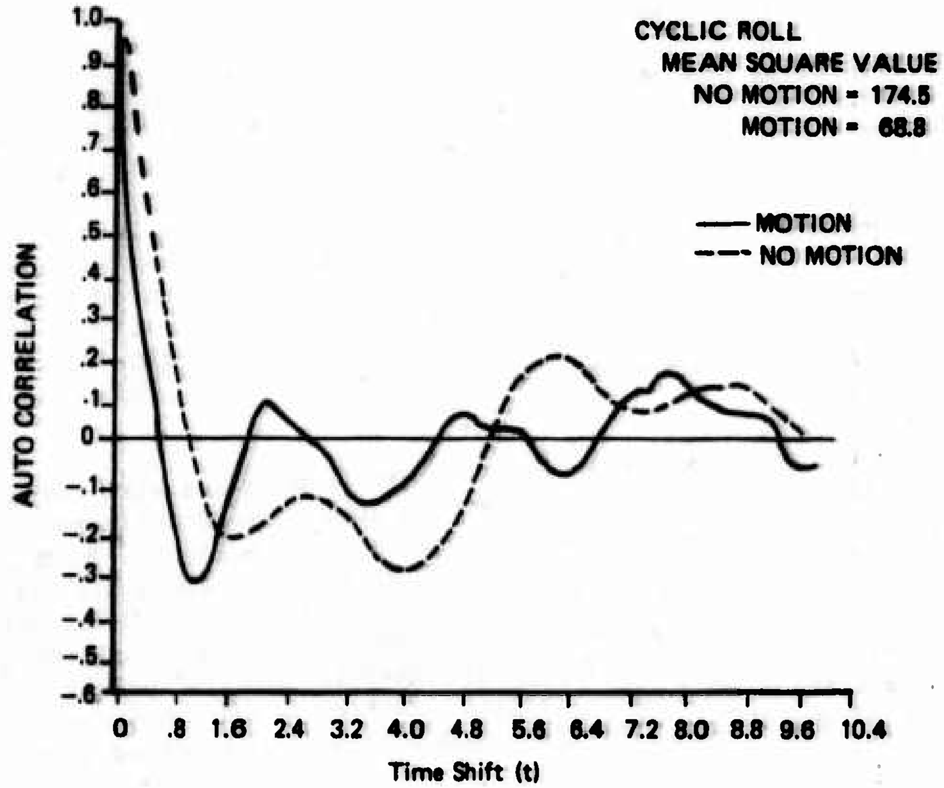
Later, in 1964 and 1965 we carried out a series of experiments using the UDOFTT simulator at NTDC in which the question of the fidelity with which the equations of motion of a high performance interceptor aircraft need be reproduced in the simulator was tested. In this series of experiments both transfer of training and control input data were collected. Conditions were tested in which the equations of motions were implemented to the fullest extent possible or were highly simplified, i.e., straight line approximations with minimal break point were used. A second variable of interest at that time was the program cycle time for updating the equations of motion. Interestingly, this cycle time problem underscores the necessity for a reliable and quick method for determining the behavioral impact of hardware questions. We need a method for quickly and effectively assessing whether or not a problem as seen by engineers really is a behavioral problem and whether or not a hardware innovation, however ingenious, is really a step forward in terms of its behavioral impact.

The results of the series of UDOFTT experiments indicated that quite drastic simplification of the equations could be instituted with no change in original learning or transfer and, more particularly, with no change in the required control inputs in order to execute control of the system in terms of the measures used to summarize the control inputs. As a footnote, very soon after these experiments were completed several quantum jumps in computer technology vastly increased the memory available, so again we trod the path of physical fidelity in implementing equations of motion rather than following any indications of the behavioral requirements which had been established, limited as they might be.

In 1973 and 1974, in an attempt to gain information with respect to the fidelity requirements for motion in ground based trainers, Matheny, Lowes, and Bynum (10) carried out a study using the TRADEC simulator at NTDC. Here again the essential question asked was "will changes in the motion cues provided to the operator change his control behavior?" In this study we were also concerned with defining the physical variable in ways which were not tied to specific pieces of equipment or methods of implementation. Rather, we wanted a general description of the motion variable which could be implemented in



Cyclic pitch output.



Cyclic roll output.

Figure 2-2. Shifts in operator control behavior different ways by design engineers according to the state-of-the-art or the constraints of cost and time. In these studies the measure of the operator's control behavior was a simple system identification method in which a Bode plot of the transfer operator's output was generated and the lower and upper corner frequencies were identified. As in former studies a shift in controller output frequencies under different conditions of the system was evident. It is of interest that this behavioral metric shows rather nicely a shift of the operator's inputs to high frequencies of an anticipatory nature when motion cues were present in this rather high frequency response system.

Incidentally, in the latter two studies, i.e., those with the UDOFTT and those with the TRADEC, the pilot ratings of the systems were obtained. The results were that there was great variability among raters with little relationship between the ratings and performance.

Most recently an attempt was made to obtain data in an instrumented aircraft which could be used to describe the operator's behavior in controlling that system under specified conditions. It was the intent that behavioral indices might be developed and used for comparing aircraft and simulator data in terms of behavioral fidelity. A word needs to be said here about the level of experience and expertise of the operator of the system when gathering data of this nature. Two points need to be made. First, the operator of the system needs to be highly experienced since he is being used as a "measuring instrument" in determining the control behavior required to control the system according to certain specified conditions. Second, from a training point of view, getting information about the control behavior of the highly experienced and proficient pilot should be very helpful in instruction or adaptive training situations since it is this type of behavior which one is trying to instill through training in the student.

In this investigation data were collected in an instrumented T-37 aircraft with experienced pilots flying prescribed maneuvers to produce system outputs within certain prescribed limits. In all, 22 variables were recorded for analysis, five of which were pilot output or control input variables. These were elevator, aileron, rudder, aileron trim tab, and throttle positions. The other 17 variables were hypothesized to be measures of inputs to the pilot. Those variables which could serve as inputs to one or more of the pilot's perception modalities were such parameters as control forces, airspeed, altitude, pitch angle, pitch rate, and so on. Similar data were scheduled to be collected in a simulator of the T-37 aircraft; however, only one pilot run was recorded due to a number of administrative and technical problems.

The method of analysis here was intended to be much more comprehensive and inclusive of more variables than any used in previously mentioned research. The analytical approach was to develop a linear, time variable, pilot identification of the pilot flying the T-37 in carrying out the specified maneuvers. A review of the methods and application of system identification techniques can be found in the April 1977 issue of Naval Research Reviews and an introduction to them is given in Brogan's (2) Modern Control Theory. Anyone seriously interested in this approach to behavior identification should read Pew et al.'s (14) March 1977 review and analysis of performance models. Due to limitations in funds and time, only a preliminary analysis of the data could be made. The analysis indicated that the data were adequate for use in carrying out an identification procedure aimed at identifying parameters and coefficients describing the operator in the aircraft. The data could also be used to compare these measures across pilots and to assess their generality. More importantly, perhaps, the data could be used to compare the results obtained in the aircraft to those from the simulator as a test of its behavioral fidelity within the limits of the identification model.

In summary, the thesis presented here is that we may establish the behavioral fidelity of a simulator for the control aspects of the task through application of system identification procedures. What is suggested is that when we speak of behavioral fidelity, we can, for the closed-loop control task, look at the operator's behavior and describe it by system identification methods. This appears to be a feasible and potentially high payoff approach to determining whether two systems are behaviorally, i.e., psychologically, equivalent. If two systems are behaviorally equivalent, we might assume that training in one would transfer positively to the other for this aspect of the task--an assumption which could be tested empirically. An advantage of such an approach is that it could be used to investigate specific characteristics of simulators both with regard to their similarity to the systems they simulate and the degree to which physical changes in them are reflected in changes in the operator's behavior.

We are very little closer to answers to the behavioral similarity question today than we were 25 years ago. It is not likely that we will get there through transfer of training experiments since these are aimed at a quite different question. Transfer experiments include variables such as the use to which the simulator is put, methods of instruction, and so forth and the operator's behavior is continually changing due to learning. In such an experiment the positive effects of a particular simulation hardware could well be offset by the negative transfer effects of method of instruction or other variables. Thus, in such experiments we have a definitive test of the transfer effectiveness of a particular training system used in a particular way but by its very nature does not give us definitive information about the fidelity of simulation questions. It seems to me we must attack this problem in a different way.

Paul Fitts, in his book Skilled Performance, seems to me to have provided a clue to this whole approach when he wrote, "in many of the tasks performed in industry and in military service an important factor is the dynamic characteristics of the physical systems such as submarines, automobiles, or other machines that are operated by man. The lags and oscillations characteristic of such systems determine part of what a man must learn in controlling them."

To me the essential ingredients for gaining some handhold on the problem are these:

1. The test of the system identification procedure for describing the behavior of the operator as he makes control inputs into the systems of interest.

2. The definition of the physical variables going into training simulator systems, and in terms which are generalizable and are not situation or hardware specific.

3. The accumulation of identification data for human operator behavior when controlling systems across representative sets of conditions.

4. In order to carry out a comprehensive program of research in this area, the availability of an instrumented system and a simulator of that system whose characteristics can be varied over a wide range, and in which important external variables can be closely controlled and made representative of those in the target system.

With this approach and these tools I believe the psychologist and the engineer could work in a symbiotic relationship providing answers as to what brings about psychological--or more precisely--behavioral fidelity.

REFERENCES

1. Benepe, O. J., Narasimhan, R., & Ellson, D. G. An Experimental Evaluation of the Application of Harmonic Analysis to the Tracking Behavior of the Human Operator. WADC Technical Report 53-384, Wright Air Development Center, 1954.

2. Brogan, W. L. Modern Control Theory. New York: Quantum Publishers, Inc., 1974.

3. Craik, A. J. W. Theory of a Human Operator and Control System. I. The Operator of an Engineering System. British General of Psychology (Gen. Sec.), 1947, 38, 56-61.

4. Ellis, N. C., Lowes, A. L., Matheny, W. G., Norman, D. A., & Wilkerson, L. E. Pilot Performance, Transfer of Training and Degree of Simulation: II. Variations in Aerodynamic Coefficients. Technical Report NAVTRADEVCEN 1889-1, Naval Training Device Center, May 1967.

5. Fitts, P. M., Bennett, W. F., & Bahrick, H. P. Application of Auto-Correlation and Cross-Correlation Analysis to the Study of Tracking Behavior. In Air Force Human Engineering, Personnel and Training Research, Air Research Technical Command, Technical Report 56-8, 1956.

6. Fitts, P. M., Bahrick, H. P., Noble, M. E., & Briggs, G. E. Skilled Performance. New York: Wiley, 1961.

7. Fulgham, D. D. Psychologists vs. Engineers. Paper presented before the Royal Aeronautical Society Flight Simulation Group, London, April 1978.

8. Hall, W. E., Jr. System Identification--An Overview. Naval Research Reviews, April 1977.

9. Matheny, W. G., Williams, A. C., Jr., Dougherty, Dora, & Hasler, S. G. The Effect of Varying Control Forces in the P-1 Trainer Upon Transfer of Training to the T-6 Aircraft. Human Resources Research Center, Air Research and Development Command, Technical Report 53-31, 1953.

10. Matheny, W. G., Lowes, A. L., & Bynum, J. A. An Experimental Investigation of the Role of Motion in Ground-Based Trainers. Technical Report NAVTRADEVCEN 71-C-0075-1, Naval Training Equipment Center, 1974.

11. Matheny, W. G. Investigation of the Performance Equivalent Method for Determining Training Simulator and Training Methods Requirements. AIAA Paper 75-108, AIAA 13th Aerospace Sciences Meeting, January 1975.

12. Matheny, W. G. Studies of Motion and Visual Interaction in Simulator Design and Application, LSI-AFOSR-FR-76-1, Final report of work under AFOSR Contract F44620-73-C-0058, September 1976.

13. Muckler, F. A., & Matheny, W. G. Transfer of Training in Tracking as a Function of Control Friction. Journal of Applied Psychology, Vol. 58, No. 5, 1954.

14. Pew, R. W., Baron, S., Feehrer, C. E., & Miller, D. C. Critical Review and Analysis of Performance Models Applicable to Man-Machine Performance Evaluation, BBN Report No. 3446, AFOSR-TR-77-0520, March 1977. (AD A038597)

15. Wilkerson, L. E., Norman, D. A., Matheny, W. G., Demaree, R. G., & Lowes, A. L. Pilot Performance, Transfer of Training and Degree of Simulation: I. Variations in Program Cycle Time and Aerodynamic Equations. Technical Report NAVTRADEVCEN 1388-2, U.S. Naval Training Device Center, 1965.

3.0 SOME CONSIDERATIONS IN DEVELOPMENT OF TEAM TRAINING DEVICES

F. Thomas Eggemeier and Bertram W. Cream
Advanced Systems Division
Air Force Human Resources Laboratory
Wright-Patterson Air Force Base, Ohio

Introduction

Several reviews of the team training area have appeared in the recent literature (e.g., Hall & Rizzo, 1975; Meister, 1976; Faust, 1976; Wagner, Hibbits, Rosenblatt, & Schulz, 1977; Collins, 1977; Kribs, Thurmond, & Mark, 1977). Each of the reviews has included an examination of current team training technology and an identification of the major shortcomings of that technology. One shortcoming consistently cited in these reviews is the absence of an adequate task analytic or Instructional Systems Development (ISD) technique for identification of team, as opposed to individual training requirements. Absence of such a technique greatly impairs the capability to optimally structure team training programs and design effective team training devices.

The purpose of this paper is to describe and discuss a task analytic technique that has been successfully used in a number of applications, including design of a team training device for members of the fire control team of the AC-130E Gunship. The technique was developed to overcome two major weaknesses of traditional ISD processes: (1) the lack of sufficient specificity for actual design of training devices; and (2) the lack of an adequate means to address design of a device for team or crew coordination training. The intent of the device design technique is to provide only the levels of fidelity that are necessary to accomplish specific training objectives. The technique is based upon careful task analyses and represents a modification and extension of traditional ISD techniques (Cream, Eggemeier, & Klein, 1978). This paper provides a general description of the technique, a discussion of the crew coordination aspects of the technique, and the results of an evaluation of the Gunship team trainer which was developed using the technique. Finally, the applicability of the technique to "man-ascendant" and "machine-ascendant" systems is discussed.

Training Device Design Technique

The training device design technique is based upon the use of behavioral data in a development process which involves the intended users of the training device, training psychologists, and simulation engineers. A basic objective of the technique is to provide a description of the training requirements that are to be accomplished in the training device. Training requirements are expressed in behavioral terms. These requirements are eventually translated into training device requirements. The user serves as a subject-matter expert in identifying the initial set of training requirements. The user also participates in the iterative process which is involved in translating training requirements into device requirements. The training psychologist is responsible for developing and coordinating inputs from the user. The psychologist also serves as the interface between the user and simulation engineer. The engineer is responsible for implementing the training requirements and producing a design specification capable of satisfying the requirements.

During initial stages of the device design process, a complete description of stimulus-response conditions for each control and display within the subject system is developed. This first analysis includes a listing of all tasks and subtasks, and an indication of their sequencing. It also includes a complete description of the initiating and terminating conditions, and the specific actions required in each circumstance.

When the behavioral task listing has been completed, the specific capabilities that will be included in the device itself are determined. The major decisions in this process are selection of the subset of tasks that will actually be trained in the device and determination of the degree of fidelity necessary to train each of the tasks in the subset.

Selection of tasks and determination of necessary fidelity is an iterative process which involves a more detailed analysis of each task. Initially, the user ranks each task on three dimensions: criticality (C), difficulty (D), and frequency (F) of performance. Tasks with high C/D/F ratings are identified for inclusion in the device, while tasks with low ratings will be excluded.

The next iteration consists of further examination of tasks identified for inclusion in the trainer. The purpose is to determine proper levels of fidelity to support training of the selected subset. During this iteration, the user provides information to identify essential stimulus and response components of the various controls and displays. For example, one console of the AC-130E Gunship trainer was a Forward Looking Infrared (FLIR) station. The FLIR operator in the Gunship is responsible for acquiring and detecting targets using infrared imagery. In design of the FLIR station, there were 26 tasks which, on the basis of C/F/D ratings, had been identified for inclusion in the trainer. It was necessary to define the degree of fidelity required to support the 26 required tasks. The visual display variables that were necessary for training were identified. For example, the users indicated that the FLIR visual stimulus used in target acquisition was a specific reflectance pattern embedded in a heterogeneous display pattern. It was also determined that this pattern was affected by a number of variables including weather, time of day, aircraft speed, approach angle, and attitude. The design team focused on the implications of not providing each of these variables, and determined that adequate training could be provided by simulating a limited number of representative cases.

Only a limited set of weather conditions and aircraft speeds was required. It was also determined that a representation of the target would be satisfactory and that all target contextual effects did not need to be duplicated.

Thus, during the iterative procedure, the critical tasks to be trained in the device were identified and the minimum level of stimulus fidelity in controls and displays necessary to support training of those tasks was determined.

Team and Crew Coordination Applications

The Gunship trainer was designed to permit both crew coordination training and individual training. It was therefore necessary to develop crew coordination training requirements in addition to training requirements that focused on the individual team members (Cream, 1974). Training consoles for four crew positions were developed: the FLIR sensor operator; and the Low Light Level TV (LLLTV) sensor operator, who detects targets using electronic imagery; the Black Crow Electric Warfare (BC/EWO) sensor operator, who protects the aircraft from enemy radar; and the Fire Control Officer (FCO), who is responsible for coordinating inputs from the sensor operators. Crew members communicate with each other verbally and also respond to system displays that provide information from other crew positions.

Because of deficiencies of traditional task analysis techniques for dealing with team/crew coordination activities, an extension of the previously described methods used for individual task requirements was developed (Cream, 1974; Cream et al., 1978). This expanded analysis was based on the results of the original C/F/D task description which had shown that over one-half of the 700 tasks identified for training involved interaction of either system, individual crew members, or both. Another crew member was therefore involved in either the initiation, action, or conclusion of the majority of individual tasks. Because of the high degree of crew coordination, a method was developed which separated the total aircraft mission into logical segments. Within each segment, the specific tasks performed by each of the fire control team members were listed. This specification included, among other items, an approximation of the time required to perform each task by each crew member. The purpose of this analysis was to: (1) describe the interactions that occurred among team members when performing coordinated actions; and (2) to gain an appreciation of the approximate task loading, not only for the duties of each team member at each flight segment, but also for actions that resulted from the combination of single tasks. Not surprisingly, it was discovered that some tasks which had originally been rated as low in difficulty or criticality when performed by a single crew member became more critical or more complex when performed within the mission context. For example, when a target is located by a sensor operator, this data is entered into the fire control system. However, before the Aircraft Commander can fire the weapons, each one of the team members has to give "permission" through both electronic and verbal communication. This apparently simple act, if disrupted or performed in an inappropriate sequence, could cause the aircraft weapons system to malfunction. The number of separate actions that each crew member was required to perform prior to "consent" varied according to the sensor system, target type, aircraft, and electronic conditions. Until the crew coordination task/timeline analysis was performed, this series of individually simple but collectively complex interactions was not apparent. There were important implications for design of the trainer. Provisions had to be made to permit these situational, context rich events to occur at both the appropriate time intervals and with the correct system displays. The requirement to provide realistic time intervals led to a further problem, i.e., the question of appropriate complexity and system fidelity versus instructional capability. Specifically, it was necessary to consider the degree to which the trainer should permit pacing of the crew performance duties according to the skill and proficiency of individual team members and the possible limitations that such pacing would have on development of appropriate experiential team skills due to trainer processing delays. A close examination of these factors and the actual tasks that were time critical or system dependent caused a modification of some design concepts that had been based originally only on the individual crew tasks. For example, modifications were necessary at the fire control officer (FCO) position. The actual FCO equipment made use of a computer to analyze target information. The trainer did not use a computer but instead required the instructor to manipulate an electronic desk calculator to compute and insert the appropriate system target information at the student FCO station. The decision to replace the computer was based on the original C/F/D analysis which indicated that this approach would permit training of critical skills with a reduction in overall trainer cost. It was apparent that this approach would cause non-real time solutions for a number of the more complex fire control problems. However, it was believed that such time differences would not cause training problems if accompanied by specific modification of relevant portions of the training scenarios. These modifications were simply to require the other trainees to perform additional related tasks during the periods when these processing time delays would be most obvious. In practice, of course, this arrangement was not as satisfactory as expected. The processing time delays were of no consequence when the instructor and the FCO student were operating in an individual mode. However, when the FCO student was required to participate as a member of the team, these delays resulted in problems for certain highly time-dependent tasks. For example, the students at the other three stations had been trained by their instructors to expect rapid responses through their equipment from the fire control system computer. Also, the students had been trained to depend on the FCO as a system integrator. Because the FCO had to wait until the instructor calculated the correct data and inserted it into the appropriate target data windows, he was obviously unable to provide his teammates with the information they needed. This inability to provide required information in a timely manner completely obviated the modified mission scenario plan and caused the performance of the FCO and other team members to deteriorate for these time-dependent tasks.

A second interesting area that was significantly affected by the crew coordination training requirement was the role of the instructor. Each position had an instructor, and there was an additional instructor who served as overall operator of the master console. Therefore, decisions regarding pacing, problem difficulty, malfunctions, emergencies, and system displays had to be coordinated among several instructors. The trainer was designed to be operated in the crew coordination mode from a set of user created scenarios (Cream, 1974). Within these scenarios, the instructor initiated problems and controlled all related activities. However, after the trainer was put into operation, it became clear that the analysis technique used had not been sensitive enough to detect a number of subtle "operational" techniques taught by experienced instructors in the team context. The specific team techniques were not part of the normal syllabus, nor were they documented within any of the technical literature. Yet, they seemed to be well known among the instructional staff.

A good example of this situation can be provided by continuation of the discussion pertaining to the FCO/time delay situation. During actual operation, the fire control team members must coordinate their target search behavior by insuring that each of the separate sensor operators is aware of the areas being searched by the other team members. This is important for two reasons. First, target acquisition often requires the coordinated action of more than one sensor. This is because different characteristics of the equipment itself make one type of target/signal more easily detected by one sensor than by another. The second reason that the team members must maintain situational awareness is to prevent being surprised by hostile fire. Both of these requirements were identified during the task analysis procedures prior to trainer construction. Some mission scenarios and target profiles were set up to exercise these skills. Because the instructors were all combat experienced, they had developed certain favorite techniques to use during target search and acquisition. These techniques involved nonserial task/control actions that served to accelerate the normal sequence of events. These time compression techniques consisted of bypassing task steps and were based upon instructor knowledge of task contribution versus the time spent in performing the step. That is, experienced instructors did not perform tasks with perceived low payoff for the particular type of target they expected to find within a specific situation. Depending upon the expected target, sensor priority, and aircraft look angle, steps within tasks were either bypassed, reordered, or reduced. These time compression techniques had not been discussed during the requirement analysis phase, nor had they been obvious during any of the flights taken by the training psychologist during task data collection. Because of the computer time delay problem at the FCO position, the student operator was unable to maintain the pace set with use of the time compression. This inability in turn prevented the instructors from demonstrating some possibly critical skills. Because of the potential consequences of this situation, special discussions were held with students and instructor personnel to develop alternate training techniques. These resulted in modifications of some mission scenarios. However, this solution was less than totally satisfactory for this small number of tasks. In spite of this modification, the trainer did perform as designed. The analysis technique did prove highly useful in identification of the major issues involved in distinguishing crew/team coordination actions from those of individual crew members.

Although consideration of such coordination requirements would appear to be straightforward, it is precisely the lack of attention to crew as opposed to individual training requirements that has been cited (e.g., Faust, 1976) as a major weakness of the ISD process. It is obvious from experience with the Gunship trainer that failure to consider team training requirements can have serious implications for design of both the crew station and instructor station. It is clear that crew team training cannot be conducted as an "add-on" to individual training in a device designed to limit fidelity levels to those required to accomplish specific individual training requirements. It is also clear that an instructor station designed only to permit control of individual skill acquisition will not permit necessary coordination among individual instructors and may not provide an opportunity for instructors to exercise unique instructional strategies that they may adopt in a team training context.

Gunship Training Device Evaluation

An evaluation of the AC-130E Gunship trainers was conducted subsequent to their design and procurement. In order to evaluate the adequacy of the trainers, the performance of crew members trained with the devices was compared to that of crew members who had not trained with the devices. Details of the evaluation have been reported previously (Cream et al., 1978). Basically, the purpose of the evaluation was to determine if crew members trained on the devices would: (1) achieve criterion performance with less airborne training time; and (2) achieve higher proficiency ratings at the completion of training than controls who were not trained with the devices. Data concerning the acceptability of the devices for training was also gathered from the users. Because of some inadvertent practice with another trainer by members of the FCO experimental group, that position was excluded from further analysis. The evaluation therefore included only the FLIR, LLLTV, and BC/EWO positions.

Results of the evaluation indicated that experimental groups reached and maintained criterion performance in significantly less airborne missions than their controls. With respect to proficiency ratings at the completion of training, the BC/EWO experimental group achieved significantly higher ratings than its control. LLLTV experimental ratings were marginally superior to controls, while FLIR experimental group ratings did not significantly differ from the controls. Overall user reaction to training devices was very favorable, as indicated by questionnaires administered to instructors and students who trained on the device.

Generally, the evaluation indicated that the trainers developed using the techniques described above were effective in providing training at each of the three Gunship positions that were analyzed. The evaluation also indicated that trainers were well accepted by the instructor and student personnel who actually used the devices.

Discussion

It is obvious from the previous discussion that the technique has been applied to training device design problems in predominantly "machine-ascendant" systems. "Machine-ascendant" systems in this context refer to those systems in which there is a high degree of man-machine interaction, in addition to whatever interpersonal interactions might occur in the system context. "Man-ascendant" systems, on the other hand, are viewed as those in which interpersonal interactions are the primary systems variables.

The distinction drawn here between "machine-ascendant" and "man-ascendant" systems closely parallels a number of other distinctions that have been drawn to characterize various approaches to team training. One of these distinctions is based on different assumptions about the nature of teams or models of team behavior (Alexander & Cooperband, 1965). The other distinction is based upon points on a continuum that can be used to characterize the task demands of the situation in which team behavior occurs (Boguslaw & Porter, 1962; Wagner et al., 1977).

One model of team behavior can be characterized as a "Stimulus-Response" (S-R) model, while the other is characterized as an "Organismic" model (Alexander & Cooperband, 1965). Team behavior in an S-R context is viewed as a series of stimulus-response units. Team training, in this view, is an attempt to apply conditioning principles and other techniques associated with individual S-R acquisition to the team (e.g., Klaus & Glaser, 1960; Glazner, 1962). The "Organismic" model, on the other hand, considers the team as a synthetic organism which is composed of individual team members (e.g., Kennedy, 1962). Development of the capability to deal with characteristics of the environment is viewed as a temporal process. Team training is therefore provided by establishing a synthetic environment in which the development process can take place.

These two models can be related to different points on a continuum of task demands under which team behavior occurs. One extreme of the continuum consists of established situations, the other extreme consists of emergent situations (Boguslaw & Porter, 1962; Wagner et al., 1977).

Established situations are basically those in which the environmental and systems conditions are specifiable and predictable, as are the probable consequences of alternative actions. Emergent situations, on the other hand, are those in which environmental conditions have not been specified, systems states do not conform to predictions, and analytic solutions to problems are not available. It has been suggested (Alexander & Cooperband, 1965) that the S-R model of team behavior most directly applies to teams operating in established situations, while the organismic model most readily applies to teams operating in emergent situations.

In considering the applicability of the current device design technique to man-ascendant and machine-ascendant systems, it should be noted that no team works exclusively in an established or in an emergent situation.

Elements of both situations appear in both man-ascendant and machine-ascendant contexts. One critical difference between man- and machine-ascendant systems may be the degree of established or emergent behavior that is required in each.

The current design technique, representing an outgrowth and modification of ISD, is certainly well equipped to deal with established situations that yield to a relatively straightforward type of S-R task analysis. Indeed, much of the technique is based on descriptions of representative stimulus configurations and required responses that are associated with the controls and displays of the particular system under study. Identification of the specific team training requirements in the Gunship trainer was based in large measure upon the structured exchange of information between operator consoles and anticipated channels of verbal communication between team members. The Gunship effort has underscored the importance of considering team training requirements in design of both the trainee and instructor stations and has demonstrated a very effective technique for doing so. Therefore, to the extent that either man- or machine-ascendant systems require behavior within established situations, the current technique is certainly applicable.

The applicability of the current technique to emergent situations, on the other hand, is less clear. Emergent situations typically involve a high element of decision making or problem solving, which have traditionally not yielded well to standard S-R task analyses. It has been noted previously (e.g., Klein, 1977; Cream et al., 1978) that not all tasks can necessarily be described as fixed sequences of discrete steps and that these tasks do not necessarily yield to S-R analysis and subanalysis. These considerations have become especially clear in dealing with highly proficient performance and with performance of highly complex tasks. Experience with the Gunship trainer is consistent with these observations. The one major area that had not been identified in the team analysis and that the Gunship trainer had not been designed to train was related to time compression techniques that were used by instructors under high task loading conditions. The time compression techniques were highly dependent upon the specific context in which the overloading occurred and represent an excellent example of emergent behavior.

It is of some interest to speculate about modifications to the current technique which might make it more sensitive to identification of team training requirements for emergent situations. One suggestion relates to a more explicit consideration of emergent situations during the training requirement identification process. Although identification of specific behaviors that relate to particular stimulus configurations would not appear feasible, it might be possible to have experienced proficient personnel identify systems parameters that might affect performance strategies, etc. Identification of these parameters would take place at a more molar level than traditional task analyses, and would have the major objective of directly specifying major systems variables that should be included in the training simulations. Such a procedure would no doubt result in higher fidelity simulations than those that result from more specific task analyses. Such higher fidelity simulations may, however, be necessary to provide the opportunity to train emergent behaviors. It is obvious that development of an analysis technique that would extend the current technique to emergent team training situations needs to be undertaken.

Another area of consideration in structuring a team training simulation for emergent situations is the instructional strategies and features that might be most profitably applied in this context. Several recent reports (e.g., Wagner et al., 1977; Kribs et al., 1977) have addressed the issue of instructional strategies within team training. This second very important area that will require continued emphasis as a strategy for team training device development is refined to reflect emergent situations.

ACKNOWLEDGMENT

Gary A. Klein provided helpful comments regarding several of the issues discussed in this paper.

REFERENCES

Alexander, L. T., & Cooperband, A. S. System Training and Research in Team Behavior. Technical Memorandum 2581, System Development Corporation, August 1965.

Boguslaw, R., & Porter, E. H. "Team Functions and Training." In Psychological Principles in System Development, Robert M. Gagne (Ed.). New York: Holt, Rinehart, and Winston, 1962.

Chapman, R. L., Kennedy, J. L., Newell, A., & Biel, W. C. The System Research Laboratory's Air Defense Experiments. Management Science, 1959, 5, 250-269.

Collins, J. J. A Study of the Potential Contributions of Small Group Behavior Research to Team Training Technology Development. Technical Report for Contract N00014-76-C-1076, Office of Naval Research, Arlington, Va., 1977.

Cream, B. W. Development of Training Requirements for a Functional Integrated Systems Trainer. Unpublished manuscript, Advanced Systems Division, Air Force Human Resources Laboratory, Wright-Patterson AFB, Ohio, 1974.

Cream, B. W., Eggemeier, F. T., & Klein, G. A. A Strategy for Development of Training Devices. Human Factors, 1978, 20, 145-158.

Faust, G. W. Team Training and ISD. In Future of Simulators in Skill Training, Proceedings of First International Learning Technology Congress and Exposition on Applied Learning Technology, Volume IV, Washington, D.C., 1976.

Glanzer, M. Experimental Study of Team Training and Team Functioning. In R. Glaser (Ed.), Training Research and Education. Pittsburgh: University of Pittsburgh Press, 1962.

Kennedy, J. L. The Systems Approach: Organizational Development. Human Factors, 1962, 4, 25-52.

Klaus, D. J., & Glaser, R. Increasing Team Proficiency Through Training: A Program of Research. Technical Report AIR 264-60-TR-137, American Institute for Research, 1960.

Kribs, H. D., Thurmond, P., & Mark, L. Computerized Collective Training for Teams. Technical Report TR-77-A4, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, Va., 1977.

Klein, G. A. Phenomenological Approach to Training. Technical Report AFHRL-TR-77-42, Advanced Systems Division, Air Force Human Resources Laboratory, Wright-Patterson AFB, Ohio, 1977.

Hall, E. R., & Rizzo, W. A. An Assessment of U.S. Navy Tactical Team Training. TAEG Report #18, Naval Training and Evaluation Group, Orlando, Fla., 1975.

Meister, D. "Team Functions," in Behavioral Foundations of System Development. New York: John Wiley & Sons, Inc., 1976.

Wagner, H., Hibbits, N., Rosenblatt, R. D., & Schulz, R. Team Training and Evaluation Strategies: State-of-the-Art. HumRRO Technical Report TR-77-1. Human Resources Research Organization, Alexandria, Va., 1977. Prepared for Defense Advanced Research Projects Agency, Arlington, Va., under ARPA Order No. 3187.

4.0 ENGAGEMENT SIMULATION TRAINING SYSTEMS: SIMULATION TRAINING IN COLLECTIVE, MAN-ASCENDANT TACTICAL ENVIRONMENTS

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

Dr. Eddowes (1978) has stated that ". . . it is axiomatic that performance on a task is facilitated if the acquisition of ability to perform the task is accomplished in a learning environment similar or identical to the environment in which performance on the task will be evaluated/measured." We can make a distinction here between "similar" and "identical" training environments. A learning environment identical to the operational environment is essentially a "replication." In a replication, there is, as Miller (1953) and Gagne (1954) have pointed out, "engineering fidelity." To achieve "engineering fidelity," the training system designer attempts to replicate the operational work environment within very tight tolerance specification, ". . . whether or not those characteristics provide critical cues or required interfaces for the trainee learning a particular job" (Miller, 1975). Replication is only possible when there is sufficient design expertise and adequate technology available to create a replica of the operational environment. It is possible, for example, to design and build simulators for certain electronic systems maintenance jobs that are replicas of the original equipment whose sole function is training, such as the Air Force's 6883 simulator that simulates the test equipment used for trouble shooting the avionics of the F-111.

When it is only possible to approximate the characteristics of the original equipment or operational environment due to inadequate design expertise or technology, then the learning environment is a "simulation" of the operational environment. The learning environment is only similar to, not identical to, the actual operational environment. For example, the visual display in most aircraft simulators is not a replica of that seen by the pilot through the windscreen of the aircraft . . . it is only a simulation and therefore an approximation of what exists in the real world.

In simulation, psychological fidelity is the critical variable. The equipment needs to resemble the original or operational system only to the degree ". . . necessary for the trainee to transfer to the job environment . . . those responses he learns to perform in the training environment" (Miller, 1975). The basic notion here is that given the trainee learns to emit a set of responses (R_{i-j}) to a set of stimuli (S_{i-j}) in a training environment, there is a set of stimuli in the operational environment (S_{k-l}) that psychologically "maps" on to the set (S_{i-j}) such that (R_{i-j}) is also elicited by (S_{k-l}). (S_{k-l}) psychologically "maps" on to the set (S_{i-j}) by occurring well within the generalization gradient of (S_{i-j}) so that there is only a minimal, insignificant discrepancy detectable in the probability that (R_{i-j}) will be emitted in the presence of (S_{k-1}), as shown in Figure 4-1. The discrepancy between P actual and P simulation can be called the "Transfer of Training Discrepancy," or TOT_Δ, for the response repertoire being trained.

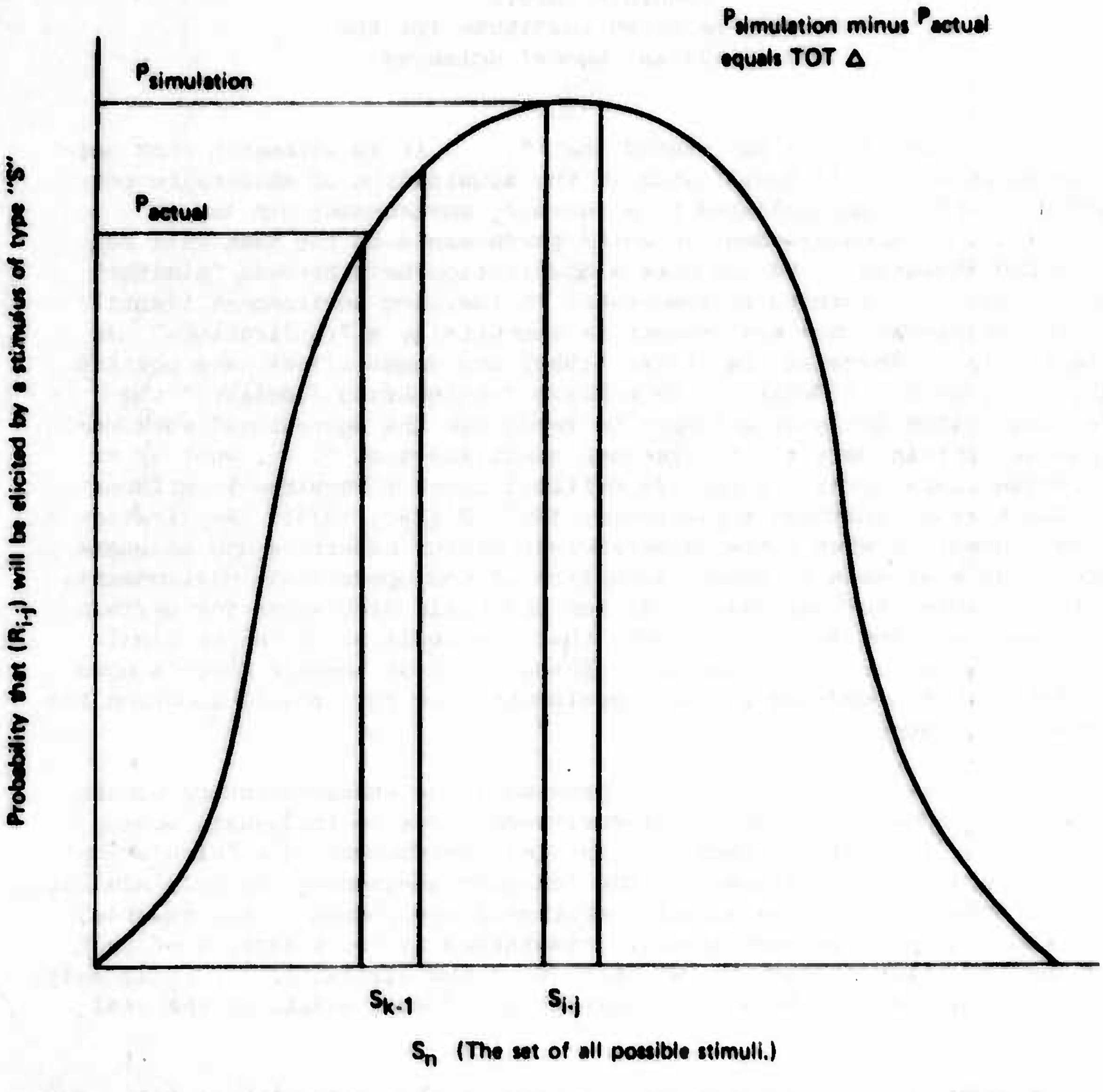


Figure 4-1. Generalization gradient for (S_{i-j}).

We can say that psychological fidelity increases when TOT_Δ decreases. TOT_Δ is an indication of the extent to which the set of actual stimuli approximates the set of stimuli that occur in the simulated environment, which are known to elicit the target response set with some probability exhibited during training. The extent of this convergence depends on a number of factors, not the least of which is individual variation in generalization gradients. But it has been possible to build training environments where the average TOT_Δ is sufficiently small and the individual variability reasonably controlled to yield a highly cost- and training-effective approach for a training population.

It is important to note that the convergence of (S_{k-l}) to (S_{i-j}), and consequently the size of TOT_{Δ}, is not necessarily a function of engineering fidelity. Eddowes (1978), for example, has found that various increases in engineering fidelity in the Advanced Simulator for Pilot Training (ASPT) have had little or no impact on certain performance variables studied. The same conclusion was reached by Martin and Waag (1978) in their work with the ASPT. Furthermore, Eddowes found that in one case, the system with a lower level of engineering fidelity yielded better in simulator performance of various flying maneuvers by inexperienced pilots.

The point here is that it does not appear unreasonable to consider psychological and engineering fidelity as different dimensions in any replication or simulation. Depending on the simulation, psychological and engineering fidelity are related by a rather complex function. Following from this assumption, increments in engineering fidelity do not necessarily decrease TOT_{Δ}. Generally speaking, then, there may be separate continua of psychological fidelity and physical fidelity in any given simulation whose precise functional relationship may vary from simple to complex, dependent upon various characteristics of the operational environment being simulated. Precisely what are these characteristics and the manner in which they affect this relationship is a matter for speculation and perhaps research. With regard to psychological or behavioral fidelity, however, there is a relatively simple first order relationship between psychological fidelity (the relative psychological contiguity of actual and simulated stimulus environments) and TOT_{Δ}.

This distinction between psychological and engineering fidelity is clearly reflected in the range of simulation training systems developed by the Department of Defense. "Representing the upper bounds of current simulation technology is the sophisticated device which reproduces, to the greatest extent possible, the complete capability and environment of a weapon system for a specific mission" (Allen, 1977). Miller (1975) calls this the "full-engineering" approach in which little or no attention ". . . is paid to the psychological factors of training equipment design. . ."; that is, little or no attention to psychological fidelity and the factors that can affect the psychological "proximity" of the target set of stimuli from the operational environment, (S_{k-l}), to those of the simulated stimuli, (S_{i-j}).

It is not difficult to understand how the full-engineering approach could overwhelm the trainee, leading to results such as those described by Eddowes (1978) and Martin and Waag (1978) for the ASPT. Any number of factors, ranging from premature loading of the attentional capacity of the trainee to improperly sequenced learning "challenges," could render the "full-engineered" simulator less effective in eliciting desired levels of performance. The solution to this sort of training problem often lies in some abstraction of the job specifically tailored to optimizing the acquisition of a limited response repertoire.

"At the lower end of the equipment capability spectrum are those devices which simulate the actual equipment only as far as external configuration . . ." (Allen, 1977). These "partial engineering" devices are designed for part-task training and may or may not include consideration of psychological factors critical for achieving psychological fidelity. Devices such as maintenance simulators and part-task flight simulators would represent this category as opposed to full mission simulators such as the ASPT.

Beyond the bounds of Allen's equipment capability spectrum is DOD's recent interest in the development of combat engagement simulation techniques. With these systems, whose objective is to simulate combat, engineering fidelity is of less concern than psychological or behavioral fidelity. The reason for this fact lies in the difference between the relative importance of the man-machine system that comprises the cockpit and fire control system of a jet fighter and that which exists in a combat unit. Emphasis on "full-engineering" solutions has been the case for primarily "machine-ascendant" systems such as flight simulators and electronic maintenance trainers in which the performance requirements are derived from the operator's interaction with the machine, via an analytic framework such as Task Analysis. Since the training objective is usually to get the operator to emit appropriate responses to stimuli emanating from the man-machine interface, the most effective reduction in TOT_Δ comes from attempting to approximate the physical characteristics of the system (that is, the machine's characteristics) as clearly as possible. But in a man-ascendant system, in which the variability of system performance is predicted in a significantly greater degree by performance in man-man interfaces, the emphasis has been naturally on psychological variables and consequently on achieving psychological fidelity in a simulated work environment.

In man-ascendant systems, critical cues and stimuli are to be found in the collective dynamics of the team or unit, particularly for those skills that involve team or unit effort and coordination. As Dr. Eggemeier (1978) has pointed out, tasks that appear to be of little importance when performed individually can become of critical importance when performed in the context of a team striving to achieve collective goals or strategies. It is the "context" of team functioning that contains the stimuli that have to be replicated or simulated in the training environment.

It is not an easy matter to analyze the "context" of team performance to isolate the critical cues and stimuli required for a high psychological fidelity simulation training system, particularly in man-ascendant systems. Man-ascendant systems pose unique analytic problems for a number of reasons. In man-ascendant systems, there is the possibility of following a systematic method for analyzing the tasks of the operators, choosing those that are critical for training, and then designing a simulator to fit the bill. Methods such as Goldstein's (1974), or the Air Force's (USAF Manual 50-2, 1970: USAF Pamphlet 50-58, 1974), allow the training system designer to isolate which displays, consoles, knobs, situations should be replicated or simulated for training the operator.

The designer can then sit down with the engineer and determine the extent to which the portion of the operational environment chosen for training can be replicated or simulated. The main stumbling block left at this point is the technology available for building the replication. Is there a CRT display big enough, or a platform motion system that can generate X g's within Y seconds in Z directions, for example. Accompanying the task analytic approach is usually an explicit description of what is to be trained, the condition under which mastery should be demonstrable, and the minimum level of acceptable performance, as Smode (1971) has described.

The ISD approach is particularly well suited to individual training situations in machine-ascendant work environments. It may also be applicable to team or crew training situations, as TAEG Report 18 (1975) argues. However, it would only seem to be applicable in certain cases in which system performance is more a function of individual-machine interaction as opposed to man-man or man-crew interaction. ". . . Tasks team members are required to perform (most notably, the team-interaction tasks) in an operational context (can) be identified and carefully analyzed," the TAEG report states. However, this approach may be somewhat simplistic, and to a certain extent erroneous, in that some of the most critical aspects or characteristics of team performance may only emerge during team interactions, the dynamics of which are not readily reducible to the framework of a task analysis. Consequently, training systems designed to teach crew or team behavior may lack those situations (the complexes of stimuli and cues) required for the elicitation of critical team skills. Alexander and Cooperband (1965) have argued that often in team training, the tasks being trained are such that exhaustive formal rules cannot be stated. Teams often develop procedures and behaviors in the process of accomplishing a task; consequently, it is often only feasible to provide teams with situations where they must innovatively and effectively develop procedures.

More recently, a number of authors have argued that for team training, the ISD-task analytic methodology does not work. Faust (1976) concludes that ". . . unfortunately, present ISD methodology as generally documented is often weak in techniques which are suitable to the identification, design, or validation of components which require team training." Faust feels, however, that innovative ISD techniques can ". . . offer considerable promise . . . if ISD is seen as a dynamic and responsive process." This opinion is somewhat more optimistic than that offered by Klein (1976) who feels that the ISD and Systems Approach to Training (SAT), as promulgated by the Air Force, have some basic weaknesses that make these approaches unsuitable for simulation design, particularly where complex individual and crew tasks are involved. Cream (1975) felt that the general family of ISD methodologies, typified by AF Manual 50-2 and AF Pamphlet 50-58, ". . . don't adequately address the complex issue of crew coordination." Cream also concludes that when system performance depends on team coordination, task analysis ". . . will not be adequate, and may possibly be misleading." Cream, Eggemeier, and

Klein (1978) have offered an alternative methodology: the Criticality, Frequency of performance, and Difficulty of performance (CFD) framework provides an interesting alternative to conventional ISD methodology, and one that is more oriented toward an effective assessment of the very complex dynamics that characterize crew and team activities.

But there are simulation design problems for crew or team activities that may not be amenable to this sort of analysis at any level. Dr. Eggemeier touched on this problem in his closing remarks in stating that the analysis of tasks and activities in team or unit-based operational environments should be terminated at a relatively global level. Cream, Eggemeier, and Klein (1975) mentioned that such behavior measured as a totality or at a global level may be too gross for training analysis or design purposes . . . but the same behavior measured as discrete steps ignores the situational or "Gestalt" qualities of the activity. Thus the dilemma.

The point being made here is that recently made by Wagner, Hibbits, Rosenblatt, and Schulz (1977), in their review of team training and evaluation strategies: there are two approaches to conceptualizing team behavior, especially for designing training strategies. First, team activities can be thought of as multi-individual activities in which individuals perform established procedures which are amenable to S-R analyses. These established procedures are explicit, probably written, and can be taught on an individual basis. For the crew to function effectively, each individual need only perform his or her tasks effectively. In this sense, the crew output represents the sum of individual activities. The second approach is that of conceptualizing the team as an organismic entity in which the activities of the individuals are mutually interdependent, and, which are defined and instigated not only in terms of one another's activities but in terms of the activity of the team, per se. Individuals working in an organismic team or crew operational environment perform tasks which are difficult to predict and analyze. These tasks or activities can be considered as emergent phenomena in that they arise from complex situations which are characterized by equally complex arrays of stimuli whose occurrence is probabilistic rather than deterministic, as far as the application of analytic techniques might be concerned. Furthermore, the response repertoires which occur in these situations can also be considered to be emergent in that the probability that the exact response will occur twice is very low due to the very low probability that the exact stimulus situation will develop repeatedly.

In such a situation, the trainee does not usually need to learn procedural aspects of the job and the operational environment, but rather the critical training requirement is usually the third phase of team training described by Kanarick, Alden, and Daniels (1971) in which teams are taught to apply their ". . . procedural and interactive skills to . . . situations requiring innovative and creative behavior." The team training requirement in this case is less that trainees need to learn specific responses and more that the individuals in the team learn to quickly and innovatively access diverse response repertoires in response to and in support of team dynamics and goals. Kanarick et al. (1971) speak of "team awareness": this is the sort of milieu or atmosphere in which is often the training objective for teams working in these sorts of unusual, complex environments.

For operational environments in which system performance is predicted primarily by the capabilities of individuals for effectively assessing diverse though appropriate response repertoires to emergent, unpredictable stimulus environments, it would seem that any sort of analysis, global or otherwise, is really not practical or cost-effective. Dr. Eggemeier's point that ". . . a global procedure which includes the C/F/D analysis could be an initial step in a solution of the fidelity specification problem in 'man-ascendant' systems" is well taken here. But analysis of the operational environment is by definition a very difficult problem when the actual occurrence of various stimulus environments is relatively unpredictable and for which there are no unique "correct" or "criterion" response repertoires. Perhaps some sort of global analysis which merely identifies the operational environment as one composed of emergent situations is the analytic tool needed!

The fidelity problem here is particularly difficult. How does the training system designer provide a set of stimuli that are designed to clearly resemble those stimuli in the operational environment when it is not possible to isolate stimuli which consistently elicit target response repertoires? And if the correct response is situation dependent, and situations are emergent, idiosyncratic phenomena, how does the designer specify behavioral objectives from which stimulus environments can theoretically be designed? And furthermore, all these machinations occur in the realm of psychological fidelity in which one of the principal training concerns is team awareness and coordination.

A very cogent example of this problem area and one in which the Army Research Institute has worked for several years is the design of training simulations for combat. Root, Knerr, Severino, and Word (1978) have discussed combat engagements as examples of complex, man-ascendant systems in which the manner and speed with which a task is accomplished is situationally determined. The situation is driven, in part, by an intelligent adversary as well as by less critical inputs such as terrain, relative force strength, and weather. Root et al. make the point that ". . . in the man-ascendant system represented by a combat unit, there can be a number of possible routes to the system's goal"; in this sort of system, the man is more a decisionmaker than an operator and often any of a number of decisions are the correct solution to the situation. Furthermore, each decision to act and act in a certain way represents an individual's response to not only a range of situational cues provided by the enemy but also a complex evaluation of behavioral options in which the individual must be aware of his role in the unit, the unit's overall mission and goal in each situation, and the unit's activities as a collective entity.

ARI has designed engagement simulation training systems for a variety of different types of combat situations involving infantry, armor, armored cavalry, air defense, and aviation. As Root et al. have discussed, it has not been possible to task analyze combat: "When considering the necessary interactions among unit elements . . . is it really practical to assume that 'all critical processes involved in successful mission accomplishment' can be identified by the task analytic process?" Consequently, the design problem has been unique, difficult, and has required a good deal of innovative effort by the staff of the Engagement Simulation Project.

In the simulation of combat, an initial assumption is made that it is not possible to specify a set of responses (R_{i-j}) for any given set of stimuli that may occur in the operational environment. That is, there is no specific response repertoire that is correct, right, or necessarily desirable for a given set of tactical stimuli. Instead, the training objective is to have units develop the capability for innovative and effective response repertoire selection to tactical situations. Root and Erwin (1977), in an attempt to operationally define high proficiency combat performance, have stated that one of the principle characteristics is effective, adaptive response to enemy dynamics, in addition to maximizing inter- and intra-unit coordinations, and maximizing friendly fire power while minimizing the effect (or potential for effect) of enemy fire power. Consequently, one of the design requirements for engagement simulation training environments is to allow sufficient freedom in the simulation for the trainee and team to choose among response repertoires, and to provide constantly varying stimulus situations so that trainees will be able to develop and practice selection of response repertoires to complex, idiosyncratic, and situational tactical stimulus environments.

In engagement simulation, this objective is realized by having two-sided, free-play combat engagements in which the "enemy" is training as well. In this way, situations unfold that are thoroughly the result of the collective dynamics of the two teams in the simulation. These situations are, as Boguslaw (1961) and Boguslaw and Porter (1962) describe, purely emergent situations: (1) all action-relevant environmental conditions cannot be predicted ahead of time; (2) the state or output of the system at any given moment is not accurately predictable with high probability; and (3) the situations are not readily amenable to analysis vis à vis the current state of analytic technology.

Although the development and prediction of these emergent situations is presently beyond the capabilities of the network of control personnel used to run engagement simulation exercises, research is being done on developing analytic frameworks for understanding the dynamics of such phenomena. Hamill (1978) and Epstein (1978) have described an Automated Tactical Operations Measurement System (ATOMS) that is designed to build a data base from engagement simulations exercises that will yield correlations between "process" activities of teams (such as movement and communication) and "product" measures (such as casualty-exchange ratios and achievement of objectives). High resolution, comprehensive data collection systems, such as ATOMS, whose analytic method is based on identifying internal consistencies within exercises that seem to be repeated between exercises, may prove to be the first step in documenting and understanding which complex behaviors in emergent situations are to be considered "good" or "desirable." This crucial first step is necessary to identify performance objectives for simulation of collective man-ascendant systems.

The present inability to analyze and document these behaviors makes it particularly difficult to develop or design an engagement simulation training system. To date, the design strategy of the Army Research Institute's Tactical Engagement Simulation Project has been to insure that certain basic design guidelines are met that have proven successful in initially designing "face-valid" simulation systems that have later been formally validated by field-experimentation. First, as mentioned earlier, an effort is made to insure that there are sufficient opportunities for teams to have to innovatively and effectively access various response repertoires as emergent stimulus environments occur. To accomplish this, care is taken that the exercise scenario or setting will support a completely free-play, two-sided, unprogramed interaction. Second, the exercise architects seek to achieve a motivational level that will drive a free-play atmosphere by building in the same behavioral contingencies that the trainee has been taught exist in combat. For example, near real-time casualty assessment appears to be one of the most critical of these contingencies. If the individual or unit knows that they can become a casualty <u>and</u> can inflict casualties, then it is possible for them to perceive the training as relevant to their survival in combat. The development of credible contingencies in the simulation motivates the troops to the point where the collective dynamics that generate emergent situations are adequately driven and fostered.

Third, for each type of exercise, weapon signatures are simulated as well as the terminal effects of various weapons by using pyrotechnics and various electromechanical devices. An attempt is made to provide the sights and sounds of battle, partly to provide the same contingencies as in combat (e.g., if the firing of your weapon system discloses your position, you can then be engaged) and partly to heighten psychological fidelity by including the violent distractors, unexpected concussions, and associated events that provide the stressful, disorganizing, attentional load that can occur in combat.

As is clearly evident, it is not possible to task analyze combat; it is not possible to specify performance objectives because there are no one-to-one correspondences between stimuli and desired responses; and it is not possible to program activities since the learning situations desired only arise from the dynamics of free-play interactions. The design goal, then, is to give trainees the opportunity to function in a combat environment, and to give the trainee the opportunity to change that environment as a function of their actions, contributing to collective dynamics.

Designers of engagement simulation training systems have evaluated the fidelity and validity of the exercises. And it is in this activity that there is a good deal of relevance to Dr. Matheny's (1978) performance equivalence model. On two occasions in 1977 and 1978, the Army Research Institute has formally validated engagement simulation training systems. These activities have been described in two reports by Banks, Hardy, Scott, Knerr, and Word (1977), and Meliza, Scott, and Epstein (1978). The purpose of these validations was to demonstrate that engagement simulation provides a training environment for exercising and learning collective tactical skills, and that engagement simulation training supports more efficient acquisition of these skills than conventional tactical training activities.

Dr. Matheny (1978) describes ". . . the performance (that) we are interested in . . . being the metric for measuring fidelity." In validating engagement simulation training systems, a very similar rationale is used. For field-validating these training systems, data are collected that document behavior which are consistently exhibited both within and among exercises. These behaviors can then be judged as similar to or disparate from those that the training system proponents would hope would be exhibited in combat. For example, proponents for engagement simulation tactical training for armor and infantry teams would expect that field exercises would provide the proper setting for eliciting effective communication and coordination between armor and infantry leaders. Furthermore, tank platoons would be expected to utilize cover and concealment techniques that follow the most recently published Army doctrine. Data collected during repeated exercises on communication rates, communicating dyads, and communication-initiated action can be used to examine whether or not "combat-like" behaviors are elicited by the simulated tactical environment. Similarly, data collected on the locations of light and heavy tank sections relative to available terrain can be used to judge whether tank platoons use cover and concealment in response to the simulated tactical situation.

Furthermore, it is possible to evaluate the criticality of various behaviors to general combat proficiency. The frequency of occurrence of behaviors and the level of proficiency at which they appear to be exhibited can be correlated with total system output, such as winning the problem, or operating at a particular casualty exchange ratio. In this way, not only can the system be validated, so to speak, but it can be evaluated by determining if it provides a stimulus environment that not only elicits correct behaviors but also provides the opportunity for development of proficiency.

Data have also been collected to determine if valid collective tactical behaviors consistently observed in engagement simulation exercises are exhibited at a higher level of proficiency in a shorter period of time than if conventional training had been used. The general finding has been that engagement simulation trains proficient collective tactical behaviors faster than conventional training techniques. The methodology that has been followed is to obtain baseline data by having two units engage in series of exercises, then having one unit train with conventional procedures while the other unit continues to conduct engagement simulation exercises, and then bringing the two units together for a final series of exercises in which performance can be compared. This procedure allows for the collection of definitive data on the training efficiency of the engagement simulation training procedures used: Does the simulation-trained team do better than the conventionally trained team in what is considered a valid test bed for collective tactical skills (i.e., the simulated combat environment)?

In summary, what are the basic characteristics of engagement simulation training systems that pose unusual problems for the training simulation designer?

1. By virtue of the operational environment being simulated, the design emphasis of the simulation has to be psychological, as opposed to physical, fidelity.

2. A combat unit is a man-ascendant system, in which both individual and collective behavior occur in emergent situations. These emergent situations are not predictable and are difficult to analyze. Consequently, it is not possible to design a simulation that yields the same stimulus environment and which elicits the same response repertoires repeatedly.

3. Engagement simulation is designed to allow for the development and occurrence of emergent stimulus environments for which the individual or unit must choose an appropriate and effective response repertoire.

4. The psychological fidelity of engagement simulation systems is measured by the extent to which the simulated environment elicits effective combat behaviors. Because it is not possible to task analyze combat due to the emergent and incredibly complex nature of unfolding events, it is necessary to show that elicited behaviors are related to total system performance, such as winning or losing a battle, or incurring "X" casualties while inflicting "Y" casualties. This demonstration, that elicited behaviors are related to total system performance, is a research and analysis problem currently being tackled by ARI staff.

Combat has proven to be a difficult challenge for the simulation-oriented training system designer. But as John Allen (1977), Deputy Director for Research and Advanced Technology in DDRE, has stated, ". . . the payoff for the military may be just as great for combat engagement simulation . . . as it is for flight simulation." And the payoff for flight simulation has been too great not to apply this approach to man-ascendant collective training problems, such as combat.

Awareness of the unique requirements facing the training system designer working with collective, man-ascendant environments is critical in building a <u>valid</u> simulation. Too often the task-analysis based ISD approach has been applied to these sorts of operational environments to the detriment of the validity and fidelity of the training simulation. The very important collective behaviors and capabilities that are elicited by and practicable in emergent situations end up not being trained. And for many of these systems, these sorts of behaviors account for the greatest portion of the system's variability. Whether it is a question of using simulation to train combat units, management teams, or elementary social skills in therapeutic or educational environments, the same unique emergent characteristics of dynamic collective systems are present. And to ignore these characteristics by using analytic techniques that cannot describe or provide design guidelines for emergent learning situations is a serious failing on the part of the training system architect.

REFERENCES

Alexander, L. T., & Cooperband, A. S. System Training and Research in Team Behavior. System Development Corporation Technical Memorandum 2851, August 1965.

Allen, J. L. The Use of Simulators: A New Thrust in the Technology Base. Defense Management Journal, 1977a, January, 30-32.

Allen, J. L. Simulation for Training in the Defense Department. National Defense, 1977b, July-August, 24-26.

Banks, J. H., Hardy, G. D., Scott, T. D., Kress, G., & Word, L. E. REALTRAIN Validation for Rifle Squads: Mission Accomplishment. U.S. Army Research Institute Research Report 1192, 1977.

Boguslaw, R. Situation Analysis and the Problem of Action. Social Problems, 1961, 3, 319-322 (as discussed in Boguslaw & Porter (1961)).

Boguslaw, R., & Porter, E. H. "Team Functions and Training." In Psychological Principles in System Development, Robert M. Gagne (Ed.). New York: Holt, Rinehart, and Winston, 1962.

Cream, B. W., Eggemeier, F. T., & Klein, G. A. Behavior Data in the Design of Air Crew Training Devices. Paper presented at the Human Factors Society Meeting, October 1975, Dallas, Tex.

Cream, B. W., Eggemeier, F. T., & Klein, G. A Strategy for the Development of Training Devices. Human Factors, 1978, 20(2), 145-158.

Eddowes, E. E. Flight Simulator Fidelity and Flying Training Effectiveness. Paper presented at the American Psychological Association Symposium entitled "Psychological Fidelity in Simulated Work Environments," August 1978, Toronto, Canada.

Eggemeier, F. T., & Cream, B. W. Some Considerations in Development of Team Training Devices. Paper presented at the American Psychological Association Symposium entitled "Psychological Fidelity in Simulated Work Environments," August 1978, Toronto, Canada.

Epstein, K. I. Automated Tactical Operations Measurement Systems: ATOMS. Paper presented at the Military Operations Research Society meeting, July 1978, Washington, D.C.

Faust, G. W. Team Training and ISD. Proceedings of the First International Learning Technology Congress: Future of Simulators in Skills Training, 1976, IV, 107.

Gagne, R. M. Training Devices and Simulators: Some Research Issues. Air Force Personnel and Training Research Center, AFPTRC-TR-54-16, May 1954.

Goldstein, I. Training: Program Development and Evaluation. Monterey, Calif.: Brooks/Cole, 1974.

Hall, E. R., & Rizzo, W. A. Focus on the Trained Man: An Assessment of U.S. Navy Tactical Team Training. Training and Analysis Evaluation Group Report, 1975, 18 (March).

Hamill, B. W. ATOMS Data Analysis. Paper presented at the Military Operations Research Society meeting, July 1978, Washington, D.C.

Kanarick, A. F., Alden, D. G., & Danick, R. W. "Decision-Making and Team Training in Complex Tactical Training Systems of the Future," in 25th Anniversary Commemorative Technical Journal, Naval Training Device Center, 1971.

Klein, G. A. Phenomenological Approach to Training. AFHRL Technical Report TR-77-42, 1977.

Klein, G. A. Problems and Opportunities in Deriving Training Requirements for the Design and Utilization of Simulators. Proceedings of the First International Learning Technology Congress: Future of Simulators in Skills Training, 1976, IV, 142.

Martin, E. L., & Waag, W. L. Contributions of Platform Motion to Simulator Training Effectiveness. AFHRL Technical Report TR-78-15, 1978.

Matheny, W. G. The Concept of Performance Equivalence in Training Systems. Paper presented at the American Psychological Association Symposium entitled "Psychological Fidelity in Simulated Work Environments," August 1978, Toronto, Canada.

Meliza, L. I., Scott, T. D., & Epstein, K. I. REALTRAIN Validation for Rifle Squads: Tactical Performance. Army Research Institute Research Report 1203, 1978, in press.

Miller, G. New Directions in Maintenance Simulation. Air Force Human Resources Laboratory (Lowry AFB), 1975, unpublished report.

Miller, R. B. Psychological Considerations in the Design of Training Equipment. Wright Air Development Technical Report WADCTR-54-563, December 1954.

Root, R. T., & Erwin, D. E. Engagement Simulation: A Training Environment for the Learning of Complex Tactical Skills. Paper presented at the Military Operations Research Society meeting, December 1976, Ft. Eustis, Va.

Root, R. T., Knerr, C., Severino, A., & Word, L. E. Tactical Engagement Simulation Training: A Method for Learning the Realities of Combat. Paper presented at the meeting of the American Educational Research Association, March 1978, San Francisco, Calif.

Smoke, A. F. Human Factors Inputs to the Training Device Design Process. Naval Training Device Center Technical Report NAVTRADEVCEN 69-C-02981, September 1971.

U.S. Air Force, Handbook for Designers of Instructional Systems, AFP 50-58, January 1974.

U.S. Air Force, Instructional System Development, AFM 50-2, December 1970.

Wagner, H., Hibbits, N., Rosenblatt, R. D., & Schulz, R. Team training and Evaluation Strategies: State of the Art. Human Resources Research Organization, Technical Report 77-1, 1977.

5.0 FLIGHT SIMULATOR FIDELITY AND FLYING TRAINING EFFECTIVENESS

Edward E. Eddowes
Flight Training Division
Air Force Human Resources Laboratory

It is axiomatic that performance on a task is facilitated if the acquisition of ability to perform the task is accomplished in a learning environment similar or identical to the environment in which performance on the task will be evaluated/measured. In the area of flying training, this compelling notion has led to an unceasing search for a learning environment for student pilots as much like the performance environment as state-of-the-art flight simulator technology can provide. While there are significant economic factors associated with the search for better flight simulator fidelity, as the business of making a simulator as much like the aircraft as possible has come to be known, this discussion will not deal with simulator costs directly. Simulator costs will be considered peripherally, later in the presentation. Instead, the focus will be on flying training effectiveness, or on the question: Does training in a high fidelity simulator transfer to an aircraft more effectively than training in a low fidelity simulator?

The facts to be presented bearing on this question will deal with a variety of flight maneuvers, learned in simulators of medium and high performance aircraft, and tested in medium and high performance aircraft. The facts were generated by both student and instructor pilots. With three exceptions, the data represent results of transfer of learning tests. In these exceptions, all data were collected in the simulator. The specific problems attacked were the training effectiveness of a state-of-the-art simulator platform motion system and three off-the-shelf visual display systems and the effects of simulator motion and visual system cueing on in-simulator pilot performance.

I. Transfer of Learning Studies

The training effectiveness of a state-of-the-art simulator platform motion system was evaluated in four different transfer of learning studies. In these studies, the Advanced Simulator for Pilot Training (ASPT) was used for the simulator pretraining. A description of the operational characteristics of the ASPT may be found in Gum, Albery, and Basinger (1975). The ASPT is equipped with a six-post synergistic motion system. In all of these studies, the motion conditions tested were full platform motion and no platform motion. The wide field-of-view visual display system of the ASPT was fully operational in each case.

The first of the studies (Woodruff et al., 1976) employed 16 Undergraduate Pilot Training (UPT) student pilots, four of whom were trained with the full platform motion system operative while another four subjects were trained in the ASPT with no platform motion. The remaining eight subjects received no ASPT training and served as a control group. Simulator pretraining covered all major areas of basic T-37 pilot training. Following each block of ASPT pretraining, all subjects were trained to Air Training Command standard criteria. Dependent variables were T-37 training hours required to reach criterion and check ride scores. Analyses of the data revealed that ASPT trained subjects required fewer hours of T-37 training and achieved check ride scores that were equal to or better than those of the control group. In addition, it was found that there were no significant or practical differences between performances of students trained under the full platform motion and the no motion conditions. The main results of these efforts are shown in Table 1.

In two subsequent studies, the contribution of a synergistic six degree-of-motion platform motion system to the acquisition of basic contact and approach and landing skills (Study I) and aerobatic skills (Study II) was reported by Martin and Waag (1978a) and Martin and Waag (1978b). Both experiments used 36 UPT students assigned randomly to three treatment groups: (1) ASPT platform motion on; (2) ASPT platform motion off; and (3) no ASPT training (control). All ASPT-trained students received the same amount of training: 10 instructional sessions in Study I, all given prior to training in the T-37, and five instructional sessions in Study II, three given before initial training in the T-37 (Aileron Roll, Split S, Loop, and Lazy 8) and two given later following three T-37 sorties (Immelman, Barrel Roll, and Cuban 8). Performance of the ASPT trained students was assessed during ASPT training by means of the ASPT automated performance measurement system (Study I) and by IP ratings and special data records (Study II). Transfer of learning was evaluated in the T-37 during the first and fifth sorties by IP ratings (Study I) and during the three initial T-37 aerobatic sorties and an additional four aircraft sorties following completion of the final ASPT training sessions (Study II).

The results of analyses of the data from these two experiments were virtually identical. There was learning during training in the ASPT for the motion and no-motion trained groups. There was significant transfer of learning to the T-37 for the ASPT trained students. The transfer effect was larger in Study I than in Study II. There were no differences in simulator training or in transfer of learning to the T-37 between motion and no-motion groups.

Another experiment evaluated the contribution of ASPT platform motion to air-to-surface weapon delivery training (Gray and Fuller, 1977). In this study, 24 recent UPT graduates were assigned equally to three groups. Two groups were trained in the ASPT, under conditions of full platform motion and no platform motion while the third group

Table 1

Average Flying Time Required To Achieve Performance Criterion

Pretraining conditions	Pre-Solo	Adv contact	Instruments	Formation	Navigation	Total
ASPT	14.2	18.7	9.0	13.7	7.1	70.7
Control	25.7	19.4	14.4	14.8	8.2	91.3
ASPT % saved	45.0	4.0	38.0	13.0	13.0	23.0

received no ASPT training. The ASPT groups were trained in 10, 15, and 20 degree weapon delivery maneuvers. Following ASPT training, the motion and no motion group subjects dropped bombs using the F-5 aircraft. Control subjects also performed the weapon delivery maneuvers in the F-5. Dependent measures were bomb miss distances and number of qualifying bombs dropped. Analysis of the data again showed no platform motion effect and significant transfer of learning from ASPT to the aircraft. It is worth noting that these subjects participated in the study following substantial T-38 training before and after completing UPT, were trained in the ASPT, a T-37 simulator, and were tested for transfer of learning in the F-5, a fighter which resembles the T-38 but in which none of the subjects had any previous flying experience. The main results of the experiment are presented in Tables 2 and 3.

Table 2

Average Bomb Delivery Circular Error

	10° dive	15° dive	30° dive
ASPT-NM	138'	144'	159'
ASPT-M	148'	138'	169'
Control	200'	180'	204'

Table 3

Qualifying Bombs

	Number	Percentage
ASPT-NM	42	46
ASPT-M	41	43
Control	24	27

The results of these four experiments indicate that simulator platform motion, which one would expect to increase the fidelity of the ASPT, had no demonstrated impact on the training effectiveness of the simulator. Although there have been criticisms of the operating characteristics of the ASPT platform motion system, and there are well recognized systematic problems with small sample studies and with the high individual performance variability characteristic of relatively inexperienced pilots, these experiments can be seen as a basis for reevaluating the previously assumed relationship between simulator fidelity and training effectiveness. For the conditions studied and the maneuvers trained, it would appear that the lower fidelity simulation was as good as the higher fidelity simulation and that the fidelity principle could be profitably reexamined, as far as platform motion systems are concerned.

Moving on now from studies of the effectiveness of simulator platform motion systems, we can review an experimental evaluation of the training effectiveness of three different types of state-of-the-art simulator visual systems (Thorpe et al., 1978). In this test of simulator visual system training effectiveness, 30 UPT graduates assigned to KC-135 Combat Crew Training (CCT) were randomly assigned to three groups for simulator pretraining using either a TV model board, a daylight color computer generated or a night-only, point light source color computer generated visual display. All students were given eight 1-hour simulator training sessions focused on the turn to final approach and landing. The transfer of learning test was accomplished in the KC-135 immediately following completion of simulator training. Additional transfer data were obtained from final checkride evaluations at the end of CCT for the 30 students in the experiment compared with the final checkride evaluations of a number of previous CCT classes. The dependent measures were IP ratings based on objective behavioral criteria.

Analyses of the data revealed that there are no differences among the three visual systems in terms of the in-simulator performance of the students. In the KC-135 transfer test, students trained using the day color and night only computer generated displays performed better than the TV model board-trained students. These differences appeared during the final approach and landing portions of the training task. In the final CCT evaluations, the day color group performed better than the night-only and TV groups, and the simulator-trained students of this study received higher ratings than previous CCT classes which had not received similar simulator pretraining.

The results of this evaluation of the training effectiveness of different simulator visual systems are interpreted as supporting the outcome of the transfer of learning tests of simulator platform motion reported above. In this case the computer generated displays were more effective than the TV-model system. The computer generated displays were less graphically realistic; that is, lower fidelity than the TV-model system. They were like cartoons of the visual environment as contrasted with a television image of a similar visual environment. While the result of the simulator visual system training effectiveness study is viewed here as showing that the less realistic, lower fidelity system was more effective, it should be noted that the hardware systems used were off-the-shelf simulator visual displays which were not matched in any systematic way on such features as resolution, field of view, or brightness but were used as intact units supplied under contract for this research effort.

To summarize this review of five transfer of learning studies of simulator training effectiveness, it becomes clear that simulator pre-training is effective, and that the differences in training effectiveness as a function of simulator realism/fidelity measured in these studies are relatively small contrasted with the transfer effects they generate.

II. Simulator Effectiveness Evaluations

This section will describe three studies in which the effectiveness of simulator design characteristics were measured in terms of the in-simulator performance of experienced pilots. The first two of these studies (Irish et al., 1977, and Irish & Buckland, 1978) investigated a relatively large number of hardware variables first in less, and then more conservative adaptations of efficient, multifactor research designs (Simon, 1973). These studies were accomplished in the ASPT and varied combinations of: (1) Platform motion conditions (no motion, three degrees-of-freedom (DOF) and six DOF); (2) G-seat, a motion cueing device which simulates sustained G-forces by inflating and deflating bladders in the pilot's seat in accordance with a pilot's aircraft control inputs (no G-seat, seat pan only, and G-seat fully operational); and (3) visual display field of view (full FOV and $36^{o}V \times 48^{o}H$, Study I, and full FOV, $36^{o}V \times 48^{o}H$, and $36^{o}V \times 144^{o}H$, Study II). Other variables were manipulated in these experiments but we will focus on the effects of platform motion, G-seat, and field-of-view variations on the in-simulator performance of eight experienced UPT IP's (three in Study I, five in Study II) who flew repeated takeoffs, overhead traffic patterns, Ground Controlled Approaches (GCA's), aileron rolls, and slow flight exercises (Study I) and overhead traffic patterns, GCA's, loops, and barrel rolls (Study II).

The results of the two studies were consistent. Performance was significantly better under wide than narrow FOV conditions. Performance was significantly better under no platform motion than under three or six DOF conditions and there was no significant G-seat effect on performance.

These studies can be interpreted as suggesting that larger sized, cartoon-style visual displays, which may be more realistic than smaller sized, cartoon-style displays, lead to improved pilot performance and therefore support the realism/fidelity principle. On the other hand, the information is not really more realistic/higher fidelity; there is simply more of it presented simultaneously. Call this a draw, as far as acceptance/rejection of the fidelity hypothesis goes.

In the case of the platform motion and G-seat variables, the pattern encountered with relatively inexperienced pilot students reappears. That is, where simulated cockpit movement and simulated G-forces are present, their absence in a simulator must be regarded as lower fidelity and the observed significantly better performance under the conditions of no platform motion and a nonoperational G-seat must indicate the ineffectiveness of these hardware features and therefore the lack of effectiveness of higher as compared with lower fidelity.

On the other hand, it must be noted that the results were the in-simulator performances of experienced pilots who are widely known to adapt rapidly so as to make the best of any aircraft control problem. So again, at best, there is only the additional suggestion that reconsideration of the validity of the assumption that simulator realism/fidelity produces training effectiveness may be warranted.

The final study bearing on the fidelity issue to be summarized in this presentation involves the use of the ASPT in a study of the effects of an Area-of-Interest display FOV size on the accuracy of air-to-surface weapon delivery performance of eight experienced tactical fighter pilots (LeMaster & Longridge, 1978).

In this study there were other variables but again we will focus only on the effects of the Area-of-Interest display FOV size. There were five different FOV sizes tested: (1) 38°V x 52°H; (2) 70°V x 70°H; (3) 70°V x 90°H; (4) 70°V x 110°H; and (5) 70°V x 130°H. The results show that bombing accuracy is significantly better when the three larger FOV sizes were used and that the bomb delivery scores obtained using these three FOV conditions are not significantly different from each other. The results are shown in Figure 5-1. The results are interpreted as a determination of the amount of realism/fidelity needed for a specific task and represent something of a novelty in simulator requirements research, a reasonably clear-cut answer to a real problem with enough excess information to apply intelligently to answer the what-if questions that often arise.

III. <u>What Does It All Mean?</u>

Q. Now with the facts before us, what can be made of them?

A. We seem to have demonstrated the ineffectiveness of a state-of-the-art platform motion system to influence the in-simulator or transfer of learning performance of UPT students in a variety of UPT training tasks. Further, the platform motion was shown to generate degraded in-simulator performance by experienced IP's in a variety of UPT maneuvers, and no consistent G-seat effect was found on the performances of the IP's in these maneuvers.

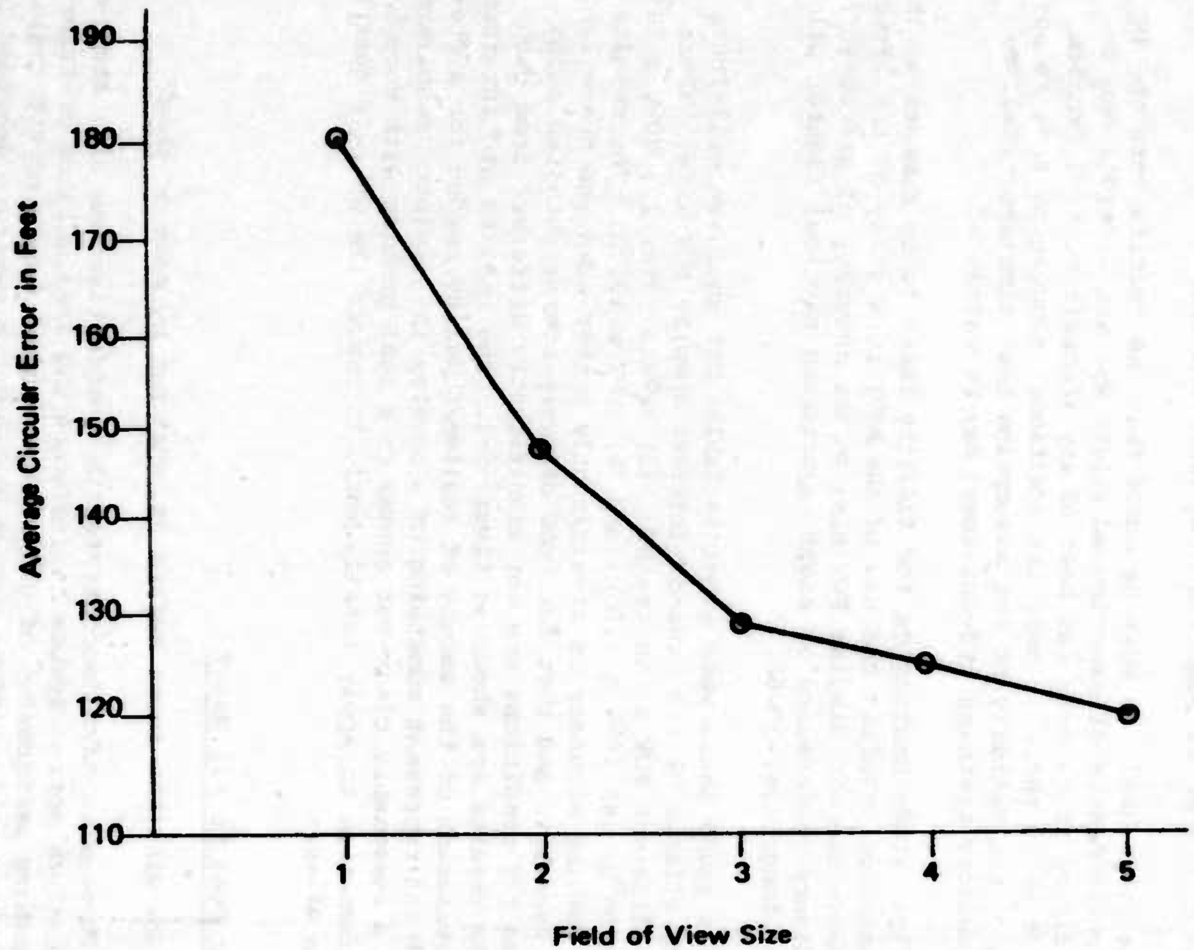


Figure 5-1. Field of view effect on simulator bomb deliveries

Q. What else?

A. The data on platform motion has been interpreted as suggesting that lower fidelity simulators may be as effective as higher fidelity simulators. This tentative conclusion is supported also, but less emphatically, by the KC-135 visual system effort. In addition, the data illustrate that the procedures used, namely the classical transfer of learning paradigm and the in-simulator differential objective performance test of skilled pilots, seem to be sensitive enough to reveal significant differences when they exist.

Q. So what?

A. To answer this question, let us consider practical issues, such as reliability, maintainability, and cost-effectiveness. It may be that these key characteristics of flight simulators are impacted by realism/fidelity requirements. Typically, higher fidelity means higher costs, lower maintainability and reliability, and, at least in some cases, as shown above, decreased cost effectiveness. Such considerations are not trivial.

Q. Finally, what shall we do about fidelity?

A. Probably it would be wise to regard it suspiciously until the research reviewed above is confirmed or disconfirmed in subsequent studies using the ASPT and other simulators in other laboratories; e.g., the Aviation Wide Angle Visual Simulator of the Naval Training Equipment Center. Meanwhile, we can remind ourselves as the design of a simulator is being determined that what can be practiced can be learned, what can be learned can be transferred, and, as soon as the capabilities of a simulator support practice of a set of training tasks, stop adding to it. Everything else the simulator does will cost more, and while it may not detract from training effectiveness, it probably won't add to it.

[illegible] Conclusions

a. The [illegible] cited [illegible] that some [illegible] in [illegible] fidelity [illegible] simulation may be as effective as [illegible] [illegible]. This [illegible] is supported also, but less [illegible], by the [illegible] system effort. In addition, the data [illegible] that the procedures that [illegible] the classical transfer of training paradigm and the [illegible] differential [illegible] performance [illegible], seem to be sensitive enough to reveal significant [illegible] than they exist.

b. [illegible]

c. To answer this question, let us consider [illegible], [illegible] reliability, maintainability, and cost-effective areas. It may be that [illegible] of flight simulators [illegible] fidelity requirements. Typically, [illegible] reliability and maintainability [illegible] in many cases [illegible] and cost effectiveness. Such considerations [illegible].

d. Finally, where should we [illegible] fidelity?

e. [illegible] research [illegible] the [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible].

REFERENCES

Gray, Thomas H., & Fuller, Robert R. Effects of Simulator Training and Platform Motion on Air-to-Surface Weapon Delivery Training. AFHRL-TR-77-29. Williams AFB, Ariz.: Flying Training Division, Air Force Human Resources Laboratory, July 1977.

Gum, Don R., Albery, William B., & Basinger, James D. Advanced Simulation in Undergraduate Pilot Training: An Overview. AFHRL-TR-75-59(1). Wright-Patterson AFB, Ohio: Advanced Systems Division, Air Force Human Resources Laboratory, December 1975.

Irish, Philip A., Grunzke, Paul M., Gray, Thomas H., & Waters, Brian K. The Effects of System and Environmental Factors Upon Experienced Pilot Performance in the Advanced Simulator for Pilot Training. AFHRL-TR-77-13. Williams AFB, Ariz.: Flying Training Division, Air Force Human Resources Laboratory, April 1977.

Irish, Philip A., & Buckland, George H. The Effects of Platform Motion, Visual and G-Seat Factors Upon Experienced Pilot Performance in the Flight Simulator. AFHRL-TR-78-9 (in press). Williams AFB, Ariz.: Air Force Human Resources Laboratory, June 1978.

LeMaster, W. Dean, & Longridge, Thomas M. Area of Interest/Field of View Research Using ASPT. AFHRL-TR-78-11 (in press). Williams AFB, Ariz.: Air Force Human Resources Laboratory, June 1978.

Martin, Elizabeth L., & Waag, Wayne L. The Contributions of Platform Motion to Simulator Training Effectiveness: Study I--Basic Contact. AFHRL-TR-78-15 (in press). Williams AFB, Ariz.: Air Force Human Resources Laboratory, August 1978.

Martin, Elizabeth L., & Waag, Wayne L. Contributions of Platform Motion to Simulator Training Effectiveness: Study II--Aerobatics. AFHRL-TR-78-52 (in press). Williams AFB, Ariz.: Air Force Human Resources Laboratory, October 1978.

Simon, Charles W. Economical Multifactor Designs for Human Factors Engineering Experiments. Technical Report No. P73-326. Culver City, Calif.: Hughes Aircraft Company, June 1973.

Thorpe, Jack A., Varney, Nicholas C., McFadden, Robert W., LeMaster, W. Dean, & Short, Lois H. The Training Effectiveness of Three Types of Visual Systems for KC-135 Flight Simulators. AFHRL-TR-78-16 (in press). Williams AFB, Ariz.: Air Force Human Resources Laboratory, July 1978.

Woodruff, Robert R., Smith, James F., Fuller, John R., & Weyer, Douglas C. Full Mission Simulation in Undergraduate Pilot Training: An Exploratory Study. AFHRL-TR-76-84. Williams AFB, Ariz.: Air Force Human Resources Laboratory, December 1976.

6.0 A DISCUSSION OF PSYCHOLOGICAL FIDELITY IN SIMULATED WORK ENVIRONMENTS

Earl A. Alluisi
Performance Assessment Laboratory
Old Dominion University

Four highly relevant, scholarly papers have been presented. Collectively, the four papers constitute an excellent state-of-the-art review of data, methods, and theory regarding the importance of functional (behavioral or psychological) fidelity versus physical (engineering or hardware) fidelity as a requirement for simulators, or for situations in which some degree of simulation is involved. Yet, the question remains unanswered, for the review indicates that the data are insufficient, the methods inadequate, and the theory undeveloped--and one of the valuable outcomes of the review provided in this symposium lies, without doubt, in the report of progress made and suggestions for directions to be followed in the future.

After presenting a carefully thought-out definition of functional (which he calls "behavioral") fidelity that differentiates it from physical fidelity, W. G. Matheny concentrates on the tracking (or "closed-loop control") aspects of flight simulation. His thesis is that system identification methods can be used to describe the operator's behavior(s) in such systems--e.g., one involving airborne flight and another involving ground-based simulated flight. If the analysis indicates that the two systems are behaviorally equivalent, it can be expected (and tested empirically) that training or practice in either would transfer positively to the other. The empirical tests are called for in order to make decisions regarding the extent to which various physical differences in alternative simulation situations reflect themselves in changes in the operator's behavior. He argues that the analysis (and test) for behavioral similarity is the crucial key to advancement toward the answer to the questions regarding the importance of different degrees or kinds of fidelity in simulation. He rejects the use of transfer-of-training data for such purposes, arguing that too many extraneous variables may influence the results to an unacceptable degree in such studies. He is pessimistic in evaluating past progress, stating that during the last quarter century little gain has been made toward answering the "behavioral similarity question." He is optimistic about the future, saying that use of the system identification procedures to study the behavioral similarity question is both feasible and of potentially high payoff.

In the second paper, F. T. Eggemeier and B. W. Cream consider the question of fidelity with respect to simulators used for crew, group, team, or unit (CGTU) training. They note that the currently available techniques of task analysis, even in the more advanced Instructional Systems Development (ISD) or Systems Approach to Training (SAT) methodologies, are inadequate for the specification of CGTU (as contrasted with individual) training requirements. They go on to propose a Criticality, Frequency, Difficulty (CFD) task-analytic technique to specify the design of training devices with "only the levels of fidelity necessary" to accomplish specific CGTU training missions. The technique involves the training psychologist as a catalyst in bringing together the user's content information and the simulation engineer's expertise. Their reported experience with the Gunship trainer clearly indicated the need to consider CGTU training requirements in the IDS/SAT process--they found substantial implications for the design of both the crew and the instructor stations. Their evaluation of the AC-130E Gunship training device showed that the trainees (a) reached and maintained criterion performance with fewer airborne missions (i.e., there was positive transfer of training), and (b) were either not different or better in proficiency than controls who were trained exclusively in airborne missions without benefit of the trainer. (Remember, Matheny would probably dispute the importance, if not the validity, of the interpretation made on the basis of any transfer-of-training study.) The authors also define, and distinguish among, (a) man-ascendant versus machine-ascendant systems on the basis of the proportion of system performance that is determined by interpersonal interactions among crewmembers, (b) S-R versus organismic models of CGTU behavior on the basis of the degree to which contextual or Gestalt-like properties do not influence outcomes in the model, and (c) the different task demands of established versus emergent situations on the basis of the extent to which desired actions (or a "school solution") can be specified. They note that the technique that they propose works best (a) with S-R analysis, (b) in established situations, (c) of either or both man- and machine-ascendant systems; its applicability in emergent situations has yet to be demonstrated.

For his paper, E. E. Eddowes adopts the view that the "goodness" or "desirability" of a given kind or amount of fidelity in a simulator should be judged in terms of the effectiveness of the device for training. He then reviews the results of five rather impressive studies of the use of flight simulation for military training. The findings clearly appear to be that pre-airborne flight experience in the way of simulator pretraining is effective for training in the sense that it reduces the requirements for airborne-flight training time. Also, within a surprisingly wide range of differences in platform motion and visual systems, the differences in simulation training effectiveness as a function of the realism or physical fidelity was surprisingly small relative to the large positive transfer effects produced by use of simulation in the first place! The practical implications of these findings--in terms of such considerations as the reliability, maintainability, and cost-effectiveness of flight simulators of different design--are quite important. And in spite of his adoption of a transfer-of-training criterion for evaluating the "goodness" of a simulator's fidelity, Eddowes approaches Matheny's position near the end of his paper when he says, ". . . what can be practiced can be learned, what can be learned can be transferred, . . .," Matheny's point earlier having been that one needs to know what has to be transferred in order to know what should be practiced and learned. Data such as those presented by Eddowes will be necessary, however, to permit the development of meaningful theory such as that called for by Matheny, and perhaps there is real justification for optimism about the future!

Such optimism is further justified by the material reviewed and presented by D. E. Erwin in the fourth and final paper. The quality or "goodness" of functional (behavioral or psychological) fidelity of a simulator or simulation situation is defined in terms of the degree of positive transfer obtained (or, inversely, in terms of the degree to which the transfer-of-training discrepancy is minimized). Functional fidelity is distinguished from physical (or engineering) fidelity; they are said to represent different (but, perhaps, sometimes correlated) dimensions, and the point is made that fully engineered fidelity may not be optimum in a simulator for training. The author then describes the approach being taken at the U.S. Army Research Institute for the Behavioral and Social Sciences in developing engagement simulation training systems--for use, of course, in collective training of CGTU's in man-ascendant tactical environments. He points out that the simulation must be to a considerable extent functional, since the physical fidelity of combat simulation is not feasible. The system simulated is man-ascendant, and the situation is emergent rather than established; the organismic model is adopted as seemingly much more representative of CGTU behavior than the S-R model. Erwin agrees with the point made previously by Eggemeier and Cream that the currently available task analytic ISD/SAT techniques appear to have little applicability in these circumstances. What is promised, however, is the development from engagement-simulation exercises of a data base of training-product measures (such as battle wins and casualty-exchange ratios) as well as potential CGTU-training-process measures (such as communication and movement patterns), with the possibility of correlating the two. The empirically valid development of such criteria would provide a major breakthrough for CGTU military training R&D; the feasibility of such a development is a major source of the optimism cited.

Some Additional Views

A discussant has the responsibility not merely to summarize or evaluate the major contributions of the presentations in a symposium, but also to add whatever additional views are judged appropriate to balance the presentation and to emphasize the state-of-affairs that seems to follow from an overview. In attempting to fulfill that responsibility, I shall slip into the first person; this should serve as a reminder that the expertise lies with the other members of the symposium--I am only a poor "outsider" looking in!

Extents of Physical and Functional Fidelity. As one of the previous speakers noted, a flight simulator does not move like the aircraft it simulates. When we speak of "physical fidelity" in a simulator, we are not speaking of a categorical characteristic that the simulator either has or fails to have. Rather, physical fidelity exists in some degree in every simulator, but to different degrees in different simulators. An underlying concern of the topic of this symposium is the question of the optimum degree of physical fidelity. Underlying, but often unstated assumptions are that increasing degrees of physical fidelity (certainly as equivalence to the system simulated is approached) are associated with (a) higher costs, and (b) increased effectiveness. Both assumptions are capable of being tested empirically, and indeed the extent to which either is valid could be demonstrated by determination of the functional relation (graph or equation) of (a) cost or (b) effectiveness as a function of degree of physical fidelity. With few exceptions, this has not been attempted for any of the simulators with which I am familiar.

On the other hand, we also speak of "functional (behavioral or psychological) fidelity." This, too, should be thought of as a continuum, rather than as a categorical characteristic of a simulator. The express concern of this symposium has to do with the optimum degree of functional fidelity. Underlying assumptions here, also often left unstated, are that relative to equal or lesser degrees of physical fidelity, functional fidelity (a) costs less for equal effectiveness, or (b) provides greater effectiveness at equal or lesser cost. As was the case with physical fidelity, these assumptions regarding functional fidelity are capable of being tested empirically. To some extent they have been, and are being, tested as indicated by the material presented in the papers of the four principal contributors to this symposium.

Although physical and functional fidelities can, and even perhaps should, be viewed as separate dimensions, they are never entirely unrelated; rather, they are something like oblique factors--they are correlated, but at best only moderately so. This being the case, there should be no surprise at finding "an engineering bias" in the direction of producing a simulator the static and dynamic characteristics of which most closely resemble those of the system simulated; i.e., a bias toward trying to design and develop and produce simulators with high degrees of physical (or engineering or hardware) fidelity.

Likewise, there should be no surprise at finding "a psychological bias" in the direction of producing a simulator that stimulates in the human operator those static and dynamic responses which most closely resemble those of the same human operator in the system simulated; i.e., a bias toward trying to design and develop and produce simulators with high degrees of functional (or behavioral or psychological) fidelity.

Simulators for Military Training. Such biases and points-of-view are seldom amenable to direct empirical verification. However, outcomes based on their application can be tested and inferences then made regarding the likely general validity of the biases themselves. That is to say, in practical problem solving situations, different biases generally lead to proposals of alternative solutions, and the solution selected as "best" serves indirectly to verify the bias out of which it was derived. The selection of one of a number of possible solutions as "best" in a practical situation generally requires a criterion, and a practical criterion generally implies some agreement on the area of use or utility of the item or solution selected.

In order to make these generalizations more concrete, let us consider the use of simulators in the Military Departments. For the most part, they are used for training or for exercising--i.e., for aiding in the acquisition of skill by essentially unskilled individuals or CGTU's, as in the use of flight simulators in undergraduate pilot training (UPT), or for aiding in the maintenance of proficiency by already skilled individuals, as in the use of flight simulators by pilots to maintain their piloting skills when for whatever reason the amount of actual flying time is deemed insufficient. Both uses require appropriate practice, with positive transfer, but there is no reason to assume a priori that what is an optimum degree of functional (or physical) fidelity for the one use will be the same as that for the other use. For the purposes of further discussion, I shall restrict myself to the use of simulators for military training. And in this regard, it might be well to think through once again why we are in the training-simulator business at all!

We start with a requirement--in this case a job we want a person to do or, from another point of view, a person who we want to be able to do a certain job. In former days, centuries ago, all craftsmen learned their trade through an apprenticeship method; even fighting men progressed to knighthood through an apprenticeship system. Years, and sometimes decades, separated the stages of skill denoted by "apprentice," "journeyman," and "master." The system worked, but it was relatively inefficient and inflexible--especially when faced with the problem of increasing quickly the numbers of available journeymen and masters. So we then developed schools and training centers to provide for increased efficiency and flexibility, especially for "apprenticeship" training such as UPT. Once we developed training centers, we had persons who began to interest themselves in the quality and efficiency of the training techniques employed. This, in turn, has led us from an "apprentice-like" system of training, where for example an instructor pilot trains a student pilot in UPT with an aircraft and no other training devices, toward a more efficient system of training that employs equipment, simulators, and training devices in addition to instructor pilots and aircraft. Now necessity, that mother of so much invention, is encouraging us to move further--to specify in concrete terms the relative merits of alternative training devices, equipments, simulators, and systems. And the concrete terms demanded are those of the cost-effectiveness of the alternatives for the purposes proposed--e.g., the cost-effectiveness of flight simulators for military training, a topic recently addressed by Orlansky and String (1977) in a two-volume report.

It seems to me that the question of the optimum degree(s) of functional (and physical) fidelity in a simulator for military training can and should be reduced to the question of the cost-effectiveness of the alternatives when used for specific and appropriate training purposes. Of course, that requires both the analysis of the cost factors, and the measurement of the effectiveness-for-training factors, for all alternatives--and for different methods of use. But these are empirical, not philosophical, questions, and although they are difficult, they are not impossible to answer.

(As an aside, this same approach should be applied at a higher level, for example to aid in making decisions regarding alternative solutions to the problem of meeting training requirements. In the training domain, we are and should be working to find ways of training people to behave in desired ways and to do so more effectively with less, and less costly, training; in the human-factors engineering domain, we are and should be working to design systems in such ways as to require less skilled persons to operate them, thereby reducing training requirements and minimizing the effects of behavioral variations on system performance. The evaluation of the training versus alternative-design-of-hardware alternatives should be on the basis of their relative cost-effectiveness ratios.)

Simulators and CGTU Training. When we move from individual to CGTU training, the problems (and our data needs) seem to increase by orders of magnitude. Perhaps because of this, until recently there has been relatively little R&D on CGTU training in the Military Departments--this in spite of the fact that the vast majority of Service training and exercising was estimated to be in the domain of collective training in units. The Army studies reported by Erwin, and the Air Force studies reported by Eggemeier and Cream are welcome indications of R&D efforts and progress in this difficult area--an area necessarily characterized at this stage with organismic models of man-ascendant systems operating in emergent situations, to which our previously developed task-analytic ISD/SAT tools seem ineffective. Yet, progress is clearly being made, and the promise for the future, e.g., in the development of both product and process criteria of CGTU training in the Army, is quite great.

The need for attention to R&D on CGTU training has been recognized and published. The Defense Science Board Task Force on Training Technology in 1976 called this "an area in which significant improvements in efficiency and effectiveness are now possible" and recommended increases in technology-base funds for training technology R&D in support of crew, group, team, and unit (CGTU) training (Alluisi, 1976, p. xi).

During that same year, the Defense Science Board Task Force on Technology Base Strategy listed "Training R&D" as the first of nine areas in which increased technology-base investment seems warranted, commenting that the development of CGTU training "shows great promise" (Rasmussen, 1976, pp. 4 & 15). In his paper today, Erwin documented John Allen's support and extension of these positions while he served as Deputy Director of Defense Research and Engineering (Research and Advanced Technology). I shall not quote him again here, but say merely that he expressed recognition of the need for increasing R&D on CGTU training, on the use of simulators in training, and on the two topics together (cf. Allen, 1977a, b).

The interests of science can be well served by the adoption of a cost-effectiveness criterion approach to the topic of this symposium and training R&D generally.

REFERENCES

Allen, J. L. The Use of Simulators: A New Thrust in the Technology Base. Defense Management Journal, 1977, January, 30-32. (a)

Allen, J. L. Simulation for Training in the Defense Department. National Defense, 1977, July-August, 24-26. (b)

Alluisi, E. A. (Chrm.). Summary Report of the Task Force on Training Technology. Washington: Department of Defense, Office of the Director of Defense Research and Engineering, Defense Science Board, February 1976.

Orlansky, J., & String, J. Cost-Effectiveness of Flight Simulators for Military Training (IDA Paper No. P-1275). Arlington, Va.: Institute for Defense Analyses, August 1977.

Rasmussen, N. (Chrm.). Report of the Defense Science Board Task Force on Technology Base Strategy. Washington, D.C.: Department of Defense, Office of the Director of Defense Research and Engineering, Defense Science Board, October 1976.