

AD-A260 971



(2)

IDA DOCUMENT D-1066

**FIDELITY AND VALIDITY IN DISTRIBUTED INTERACTIVE SIMULATION:
QUESTIONS AND ANSWERS**

Norman E. Lane
Earl A. Alluisi

November 1992

93-03939



DTIC
ELECTE
FEB 25 1993
S E D

Prepared for
Defense Advanced Research Projects Agency

Approved for public release; distribution unlimited.



INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

98 2 24 061

DEFINITIONS

IDA publishes the following documents to report the results of its work.

Reports

Reports are the most authoritative and most carefully considered products IDA publishes. They normally embody results of major projects which (a) have a direct bearing on decisions affecting major programs, (b) address issues of significant concern to the Executive Branch, the Congress and/or the public, or (c) address issues that have significant economic implications. IDA Reports are reviewed by outside panels of experts to ensure their high quality and relevance to the problems studied, and they are released by the President of IDA.

Group Reports

Group Reports record the findings and results of IDA established working groups and panels composed of senior individuals addressing major issues which otherwise would be the subject of an IDA Report. IDA Group Reports are reviewed by the senior individuals responsible for the project and others as selected by IDA to ensure their high quality and relevance to the problems studied, and are released by the President of IDA.

Papers

Papers, also authoritative and carefully considered products of IDA, address studies that are narrower in scope than those covered in Reports. IDA Papers are reviewed to ensure that they meet the high standards expected of refereed papers in professional journals or formal Agency reports.

Documents

IDA Documents are used for the convenience of the sponsors or the analysts (a) to record substantive work done in quick reaction studies, (b) to record the proceedings of conferences and meetings, (c) to make available preliminary and tentative results of analyses, (d) to record data developed in the course of an investigation, or (e) to forward information that is essentially unanalyzed and unevaluated. The review of IDA Documents is suited to their content and intended use.

The work reported in this document was conducted under contract MDA 903 89 C 0003 for the Department of Defense. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE November 1992	3. REPORT TYPE AND DATES COVERED Final--November 1991-November 1992
4. TITLE AND SUBTITLE Fidelity and Validity in Distributed Interactive Simulation: Questions and Answers			5. FUNDING NUMBERS C - MDA 903 89 C 0003 T- A-132
6. AUTHOR(S) Norman E. Lane, Earl A. Alluisi			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses 1801 N. Beauregard St. Alexandria, VA 22311-1772			8. PERFORMING ORGANIZATION REPORT NUMBER IDA Document D-1066
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Defense Advanced Research Projects Agency 3701 N. Fairfax Drive Arlington, VA 22203-1714			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) Distributed Interactive Simulation (DIS) involves the on-line networking of large numbers of participants operating through simulators, actual equipment, or computer models of friendly and opposing forces in free-play exercises on a simulated battlefield. DIS employs new simulation-related technologies that can be used to support innovative applications generally not feasible with the older technologies of conventional simulation. It has the potential of revolutionizing future work in (a) collective training (b) the development and evaluation of tactical concepts and doctrine, (c) system test and evaluation, and (d) weapon system concept analysis. Its potential importance is sufficiently great to suggest that special attention be given to issues regarding the desirability or necessity of relevant simulator characteristics. Two such issues are addressed in this document. Specifically, in this document, simulator <i>fidelity and validity</i> issues and applications-based fidelity requirements are addressed in a question-and-answer dialogue format. Secondly, four key drivers of fidelity requirements are identified: (a) the mission to be simulated, (b) the objective(s) of the simulation; (c) the fidelity dimensions, and (d) the simulation components. Thirdly, some conclusions about DIS and the associated fidelity issues are discussed. It is suggested that two parallel trends--the emergence of affordable technology and the continuing "drawdowns" in military resources--will interact to increase the importance of DIS as a problem-solving tool across a wide variety of potential application areas. It is suggested that the advantages and disadvantages of such DIS applications should be examined carefully and logically for they are capable of providing a powerful supplement or alternative to large-scale field exercises.			
14. SUBJECT TERMS distributed interactive simulation, simulation, training, test and evaluation, simulator fidelity, simulator validity			15. NUMBER OF PAGES 37
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR

IDA DOCUMENT D-1066

FIDELITY AND VALIDITY IN DISTRIBUTED INTERACTIVE SIMULATION:
QUESTIONS AND ANSWERS

Norman E. Lane
Earl A. Alluisi

November 1992

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

DTIC QUALITY INSPECTED 3

Approved for public release; distribution unlimited.



INSTITUTE FOR DEFENSE ANALYSES

Contract MDA 903 89 C 0003
DARPA Assignment A-132

ACKNOWLEDGMENTS

The authors acknowledge with thanks the helpful comments, suggestions and contributions of many persons who helped in the development and presentation of the concepts contained in this document. Ms. Lee Ann Miller provided outstanding support in preparation of the manuscript. Commander Dennis McBride furnished both encouragement and penetrating (but constructive) criticism throughout. Drs. J. Dexter Fletcher, Jesse Orlansky, Joel Schoen, and Henry L. Taylor each made recommendations for the improvement of earlier versions, and most, but not all, of their advice has been incorporated into this final edition. However, since they made all final decisions regarding the paper, the authors must claim sole responsibility for all errors and faults that remain.

ABSTRACT

Distributed Interactive Simulation (DIS) involves the on-line networking of large numbers of participants operating through simulators, actual equipment, or computer models of friendly and opposing forces in free-play exercises on a simulated battlefield. DIS employs new simulation-related technologies that can be used to support innovative applications generally not feasible with the older technologies of conventional simulation. It has the potential of revolutionizing future work in (a) collective training, (b) the development and evaluation of tactical concepts and doctrine, (c) system test and evaluation, and (d) weapon system concept analysis. Its potential importance is sufficiently great to suggest that special attention be given to issues regarding the desirability or necessity of relevant simulator characteristics. Two such issues are addressed in this document.

Specifically, in this document, simulator *fidelity and validity* issues and applications-based fidelity requirements are addressed in a question-and-answer dialogue format. First, some basic terminology and definitions are presented. Attention is focused on the imprecision of the term *fidelity* and its limited utility in describing complex simulations (some 22 different definitions of fidelity have been used in various contexts). Some alternative ways of viewing fidelity concepts are presented.

Secondly, four key drivers of fidelity requirements are identified: (a) the mission to be simulated, (b) the objective(s) of the simulation; (c) the fidelity dimensions, and (d) the simulation components. Based on these drivers and on the premise that each decision about the configuration of a simulation must be justified (or "anchored") by a defined simulation requirement, the concept of *fidelity anchoring* is introduced as a way of systematically making decisions about fidelity requirements in DIS.

Thirdly, some conclusions about DIS and the associated fidelity issues are discussed. It is suggested that two parallel trends--the emergence of affordable technology and the continuing "drawdowns" in military resources--will interact to increase the importance of DIS as a problem-solving tool across a wide variety of potential application areas. It is suggested that the advantages and disadvantages of such DIS applications should be examined carefully and logically, for they are capable of providing a powerful supplement or alternative to large-scale field exercises.

CONTENTS

Acknowledgments	ii
Abstract	iii
Figures	v
Abbreviations	vi
INTRODUCTION	1
QUESTIONS AND ANSWERS	4
1. What is different about <i>Distributed Interactive Simulation</i> (DIS) versus more traditional simulation? Why does that matter?	4
2. What is meant by <i>fidelity</i> in simulation?	5
3. How do fidelity requirements differ across different applications of DIS?	6
4. I understand that fidelity is a complicated term. Is there a better way to talk about all these different aspects of fidelity?	8
5. But don't I still need to know how much fidelity, realism, validity (or whatever) I should have for a particular simulation?	9
6. Is it really possible to have too much fidelity?	10
7. When I ask about fidelity, people keep telling me, "It all depends." Why can't I get a straight answer?	11
8. What was that about <i>fidelity anchoring</i> ?	11
9. OK, I understand now what fidelity anchoring is intended to do. How does one go about doing that?	12
10. So we have four sets of dimensions. How do we use them to make fidelity decisions?	17
11. I'm confused about the logic of using the <i>highest</i> fidelity requirement. Aren't we trying to <i>reduce</i> unnecessary fidelity?	20
12. We've spent a lot of time on fidelity anchoring. Why are you so concerned about it?	20
13. I notice that fidelity anchoring uses relative judgments of fidelity requirements (high, medium, or low). Isn't there some better way to quantify fidelity?	21
14. That's enough about fidelity for now. Let's backtrack a bit. You keep making a distinction between basic skills and warfighting skills. What's the difference and why does it matter?	23

15. If fidelity requirements are only dependent on intended use, why do people worry so much about <i>negative training</i> ?	24
16. You mentioned earlier the <i>validity</i> of a simulation as another way of talking about fidelity. Can you expand on that idea?	25
17. OK, that just about wraps it up. Any last comments?	26
References	27

FIGURES

1. A Representative Mission Profile	13
2. Simulation Objectives (Activities To Be Trained, Practiced, or Evaluated)	15
3. Fidelity Dimensions	16
4. Components of a Distributed Interactive Simulation (DIS)	18

ABBREVIATIONS

DIS	Distributed Interactive Simulation
OPTEMPO	Operating Time
IST	Institute for Simulation and Training
SIMNET	Simulation Network
T&E	Test and Evaluation
WCA	Weapons Concept Analysis
TER	Transfer Effectiveness Ratio

INTRODUCTION

Distributed Interactive Simulation (DIS) involves the on-line networking of large numbers of participants--simulators, actual equipment and computer models of friendly and opposing forces--in free-play exercises on a simulated battlefield. The simulation is *distributed* in that players may be geographically widely separated. It is *interactive* in that what each participant does may have immediate consequences for other participants, as well as for himself.

The development and successful implementation of DIS exercises is supported by the emergence of a host of new simulation-related technologies; taken together, these technologies enable uses of simulation across a variety of application areas not feasible with conventional simulation approaches. Because these applications of DIS are so sharply contrasted with previous simulation approaches, there is sometimes misunderstanding or confusion about how the technology works and what it is intended to accomplish. A major area of confusion has to do with the nature of *fidelity* and the degree of fidelity or validity required in applications such as warfighting training and weapons concept analysis.

Some of these issues are addressed in this paper in the form of an imaginary dialogue between (a) an informed, mildly skeptical observer with interests, but no particular expertise, in DIS, and (b) an advocate of simulation as a problem-solving technique who is especially interested in DIS technology and its applications. The imaginary dialogue outlines some current and future characteristics of DIS, and covers some basic terminology and definitions, particularly those having to do with problems in the use of a concept of fidelity that is basically imprecise and therefore not a very useful term to employ with reference to simulations.

In addition to discussions of fidelity terminology, basic skills and warfighting skills are distinguished, and the importance of their differences for DIS applications is underscored. Relations between the companion concepts of fidelity and validity are reviewed as they apply to DIS simulations. Four key dimensions or *drivers* of fidelity requirements are identified and their interactions examined: (a) the mission to be simulated, (b) objective(s) of the simulation, (c) fidelity dimensions involved in the simulation, and (d) simulation components to be represented in the simulation. Based on these drivers, a

systematic method for making decisions about fidelity requirements in DIS, called *fidelity anchoring*, is introduced, and some of its advantages are discussed. Several possible classes of applications are suggested to illustrate the broad potentials of DIS, and the differing fidelity requirements of each class are analyzed. The paper concludes with the contention that DIS will increase in importance as a problem-solving tool as "drawdowns" continue in military resources. The only alternative to DIS for certain kinds of training and analysis is the use of large-scale field exercises, the conditions of which are increasingly more difficult to meet. Thus, while neither a panacea nor a threat to field exercises or *operating time* ("OPTEMPO"), DIS is an alternative whose advantages and disadvantages should be considered logically and rationally. At minimum, the use of DIS is clearly a means of increasing the effectiveness and efficiency of whatever OPTEMPO is available.

The imaginary dialogue is in the form of 17 questions and answers. The questions are:

1. What is different about distributed interactive simulation (DIS) versus more traditional simulation? Why does that matter?
2. What is meant by *fidelity* in simulation?
3. How do fidelity requirements differ across different applications of DIS?
4. I understand that fidelity is a complicated term. Is there a better way to talk about all these different aspects of fidelity?
5. But don't I still need to know how much fidelity, realism, validity (or whatever) I should have for a particular simulation?
6. Is it really possible to have too much fidelity?
7. When I ask about fidelity, people keep telling me, "It all depends." Why can't I get a straight answer?
8. What was that about *fidelity anchoring*?
9. OK, I understand what fidelity anchoring is intended to do. How does one go about doing that?
10. So we have four sets of dimensions. How do we use them to make fidelity decisions?
11. I'm confused about the logic of using the *highest* fidelity requirement. Aren't we trying to *reduce* unnecessary fidelity?
12. We've spent a lot of time on fidelity anchoring. Why are you so concerned about it?

13. I notice that fidelity anchoring uses relative judgments of fidelity requirements (high, medium, or low). Isn't there some better way to quantify fidelity?
14. That's enough about fidelity for now. Let's backtrack a bit. You keep making a distinction between *basic* skills and *warfighting* skills. What's the difference and why does it matter?
15. If fidelity requirements are only dependent on intended use, why do people worry so much about *negative training*?
16. You mentioned earlier the *validity* of a simulation as another way of talking about fidelity. Can you expand on that idea?
17. OK, that just about wraps it up. Any final comments?

The reference materials that were used in the development of the questions and answers are listed at the end of the paper, and they are also linked, where appropriate, to the separate questions.

QUESTIONS AND ANSWERS

1. What is different about *Distributed Interactive Simulation* (DIS) versus more traditional simulation? Why does that matter?

Distributed Interactive Simulation (DIS) involves networking or linking together, in real time, some number of manned simulators with each other, and sometimes with actual equipment or computer models. Each of these participants--simulators, operational equipment, or models--is called an *entity*, and each entity uses its own (and probably different) computer system to put data on and take data off a simulation or communications network. These data describe in real time the changes that would take place in what each entity would see, hear, or otherwise sense if the simulated battlefield activities were real. The simulation is *interactive*, so actions taken by one entity may have direct consequences for other entities, and are immediately reflected in changes to what other entities can see, hear, or sense. Battlefield situations are changed dynamically by entity activities, and because of this free play, battle outcomes are never predeterminable, but rather evolve as a direct result of interactions among the players in the simulation.

DIS requires that the simulations be *interoperable*. Thus, simulators and equipment at varying levels of sophistication and fidelity must be able to participate in the battle on equal terms, even if they are built with different computer architectures. There must be a "level playing field." Because each individual entity may use a different way of coding or representing relevant information, special attention has to be given to combining these many different formats into a single standard set of communication protocols by which *data packets* are passed around the network. Developing and maintaining these protocols and other requirements of interoperability add to DIS significant elements of complexity not present in the typically conventional simulations.

The simulation is *distributed*, so entities may be widely separated geographically--even worldwide. Under such conditions, successful operations require careful considerations of communication delays to maintain realistic real-time interactions. Current networks could involve a thousand or more entities, each of which would have to examine all individual packets of information on the network to determine the specific self-relevance

of each. Future DIS networks might have to link hundreds of thousands of entities, and thereby generate so many packets that no single participant could examine and respond in real time to all of them. Special processors such as *intelligent gateways* will be needed for these large future DISs. Each gateway would service a large group of entities, monitoring the network information traffic, sorting out and forwarding what each entity needs to know based on its unique requirements. Development and use of such intelligent gateways will enable almost unlimited participation in battlefield simulations by entities throughout the world.

Papers suggested for further reading on topics related to Question 1 include Alluisi (1991), Institute for Simulation and Training [IST] (1990, 1991), Repolewski (1990), and Thorpe (1987).

2. What is meant by *fidelity* in simulation?

People tend to use the term fidelity as a kind of shorthand for describing how closely a simulation corresponds to the "real thing." We can talk about a *reference situation*, in which the real system is performing a specific mission under specific operational conditions. If we could define the reference situation precisely, we could think of fidelity conceptually as the degree of correspondence between our simulated situation and the reference situation. Unfortunately, there are just too many ways in which missions and conditions can vary for this sort of conceptual definition to be of much help in describing the fidelity of any particular simulation. In fact, unless we add a great many additional modifiers, the term fidelity is so general as to be almost meaningless in asking simulation questions. For example:

There is *no single definition* of fidelity. Attempts to make the term less vague have caused distinctions to proliferate; at least *22 different definitions* have been used in the literature to refer to different kinds of fidelity (physical, equipment, psychological, perceptual, functional, procedural, task, logistic, threat, etc.). Each of these definitions could be appropriate in some application.

A simulation can be subdivided into as many as *20 different components* (workstation, visual display, controls, data base, etc.). Each of these components may have a *different level* of fidelity associated with it, depending on the intended use of the simulation system.

Some authorities think that the term, fidelity, should apply only to the *hardware* (does it look like and operate like the actual equipment). Others think that it should include some or all of the *other components and representations* in the complete simulation system, including data bases

(threats, terrain), communications and information flows, physical and sensory environments (motion, noise, heat/cold), and so forth.

Simulation, particularly DIS, can be used for many different purposes. Although it is particularly valuable for training what we call *warfighting skills*, as in the SIMNET applications, DIS can also be used for design, test and evaluation of new systems, or for looking at how systems might best be used (new tactics, concepts, and doctrine). A particularly valuable use is for *weapons concept analysis*, in which warfighting scenarios can be played out with a variety of potential engineering or manpower solutions to estimate the battlefield impact of implementing a proposed solution. Each of these different purposes--warfighting training, test and evaluation, tactics development, or analysis--emphasizes different aspects of the general fidelity concept, and therefore suggests a different measure (and weighting for importance) of fidelity and fidelity issues.

Notice that by defining the problems associated with fidelity as a descriptive term with limited meaning, we have created something of a problem for ourselves in future discussions. There is as yet no widely accepted lexicon of terms that conveys adequately the multiple connotations implied by common uses of the term, fidelity. We do not wish at this point to propose any additional terminology to further confound an already cloudy definitional problem. We will thus continue, albeit reluctantly, to use *fidelity* in its broadest and most casual sense, as a shorthand reference for the overall agreement between a simulation and some general conception of operational reality.

Papers suggested for further reading on topics related to Question 2 include AGARD (1980), Blaiwes and Regan (1986), Cream, Eggemeier, and Klein (1985), Hays (1981), Hays and Singer (1989, Chs. 1-3), Kinkade and Wheaton (1972), Meister (1990), Miller (1954), Rankin, Bolton, Shikiar, and Saari (1984), Semple, Hennessy, Sanders, Cross, Beith, and McCauley (1981), and Su (1984).

3. How do fidelity requirements differ across different applications of DIS?

We noted four broad classes of DIS use--(a) system test and evaluation, (b) training in warfighting skills, (c) the development of new tactics or concepts and doctrine, and (d) weapons concept analysis. Although all of these are important applications of DIS, the fidelity questions in each are distinct. Because the intended uses of simulation are different, and because different components are emphasized, the requirements for fidelity have to be considered separately for each of the classes of intended use.

Test and evaluation (T&E) is directed toward finding out if the equipment can perform "as advertised"--that is, if the hardware meets its design objectives. The focus of attention is on the *system hardware* and all its subsystem-hardware components as they interact in a realistic environment. Because the system is highly evolved by that stage, the equipment, mission, and environment will normally be represented in considerable detail. It is important that the *outputs* be precise enough to match the detail of the design objectives, and sufficiently diagnostic to permit identification of specific system deficiencies, where they occur. The models and simulations employed for T&E purposes will normally have for all the represented components the highest levels of fidelity that technology and resources permit.

In the *training of warfighting skills*, the focus of attention is on the human components more than on the hardware. The objective is to provide operators with *practice* in using their platforms, weapons, and related equipment as *weapons systems* operating in a coordinated battle situation in response to a common threat. For each of its components, the simulation requires only that level of fidelity which allows realistic practice to occur and enables the desired set of training objectives to be met. It is important, for example, that the simulation allows operators to obtain correct cues, to follow their standard procedures, and to receive appropriate feedback. On the other hand, it is *not* important that the simulation present a perfect functional replica of the actual equipment. Only those components that are needed for the *intended training use* should be represented in full detail and with "full fidelity." Because warfighting training involves the learning and practice of complex procedural and cognitive skills, the simulation focus needs to be on precision in replicating *operator-controlled processes*, more than on the accuracy of hardware representation.

In *tactics or concepts and doctrine development*, the simulation is employed with men-in-the-loop, but with equal emphasis on the hardware (weapon systems) and the humans who operate and maintain the systems. This usage has to do with the establishment of the principles of warfare, as well as the procedures, as they are to be practiced by our forces. In short, it is to try out new concepts and doctrine, to identify the optimum alternatives, to test their generality in different simulated battlefield situations, and finally *to establish the new concepts and doctrine* in the Services' documentation for operations and training. Since the purpose of this type of simulation is to examine the possibility of unanticipated events, an approximate, but not necessarily precise, level of fidelity is acceptable.

Weapons concept analysis (WCA) involves the evaluation of *ideas for new weapon capabilities*--ideas for what are sometimes called *notional systems*. Given the availability of a DIS that represents some battlefield, such an analysis can be made even if only a preliminary outline is available regarding how the notional system might function. The objective of the WCA is to determine if and how the new capability might affect battlefield outcomes, over and above the outcomes attainable with existing capabilities (i.e., the marginal contribution of the notional system). Based on those outcomes, the decision can be made on whether further development of the concept is warranted, and if so, to define the ranges of key parameters to be associated with the notional system. WCA is a relatively coarse screen; it is intended to permit early elimination of those concepts that lack merit as potential solutions to identified problems, and on the other hand, to highlight the concepts that have the greatest potential payoff. Thus, the simulations that support WCA need *not* be extremely precise, since the objective is to sort concepts into broad decision categories (e.g., *to discard, to withhold judgment until more data are available, or to accelerate development*). Likewise, the fidelity required for such decisions is somewhat lower than that needed for most of the other modeling and simulation applications.

Papers suggested for further reading on topics related to Question 3 include Hays and Singer (1989, Chs. 1, 12), Hodges and Dewar (1992), and Meister (1990).

4. I understand that fidelity is a complicated term. Is there a better way to talk about all these different aspects of fidelity?

In the previous response to Question 2, we indicated that there is *no single definition* of fidelity, but rather that at least *22 different definitions* have been used in the literature to refer to different kinds of fidelity (physical, equipment, psychological, perceptual, functional, procedural, task, logistic, threat, etc.). Although not everyone agrees with all the details, there are now some appealing trends toward clustering these definitions into three different terms (*fidelity, realism, and validity*) to represent or take the place of both the older idea of fidelity being a unitary property or the later identification of more than a score of different possible fidelity properties of simulations.

In the new usage, fidelity is strictly an *engineering term* that refers only to the *physical correspondence* of the simulator's hardware to that of the actual equipment being simulated. Do they look the same? Do the controls and switches work the same way?

The second concept is that of *realism*. *Realism* refers to the *perceptions and subjective judgments* of the people using the simulations--do they perform and appear

sufficiently close to the real systems and equipments (to reality, in brief) to permit and support effective training or evaluation? To the extent that the answer to this question is yes, the simulation can be said to exhibit an associated degree of realism.

Validity is the third of the three concepts. *Validity* refers to the *suitability* of the simulation for a *specific application*. Does the simulation permit the operator to attain a desired training objective or level of task proficiency? Can a new system or tactic be evaluated at a level of precision that would permit or support decisions regarding the system's probable effectiveness? (*Validity* also has some different specific meanings for the military modeling and simulation community; more will be presented on these differences later in this paper.)

Using these three terms allows for a much clearer examination of the issues involved in asking simulation questions. For example, fidelity (engineering) as such is not very important for simulation except as it impacts validity, although it is (or can be) a major cost driver. *Realism* is important largely because of its role in affecting user acceptance; indeed, sometimes simulations that vary considerably in engineering fidelity differ only unimportantly in perceived realism. *Validity* is important because it reminds us that any judgment of the suitability of a simulation can be made only in reference to its intended use in a *specific and well-defined application*.

The word fidelity may sometimes be acceptable as a convenient shorthand for conveying a general idea of likeness between the simulation and the "real world." Use of the word in that way, however, sheds little or no light on the underlying issues (i.e., issues of fidelity, realism, or validity--engineering, perception, or suitability). Rather, it tends to get in the way of communicating about specific simulations, and confuses the more precise usages that are necessary to address the simulation issues of greater importance. In the long term, new and more definitive terminology must be developed to support discussion of these issues.

Papers suggested for further reading on topics related to Question 4 include Hodges and Dewar (1992), Jones, Hennessy, and Deutsch (1985), and Williams and Sikora (1991).

5. But don't I still need to know how much fidelity, realism, validity (or whatever) I should have for a particular simulation?

The obvious answer would seem to be, "All you can afford." That is probably correct for some applications like test and evaluation, but it is often not true for training.

Sometimes all the fidelity you can afford isn't good enough, so you have to use real equipment in real-world field exercises even though such operating time (OPTEMPO) is more expensive. For example, the finer points of vehicular control cannot be fully trained with either DIS or conventional simulation; practice in controlling actual vehicles is required to attain the higher levels of motor-control skills.

At other times, all the fidelity you can afford may be *too much* for optimum training--for example, at early stages in the learning process when the trainee is likely to be confused by the full complexity of the weapon system he is trying to learn to operate or maintain. The trainee is then more likely to learn better and faster if presented with simplified representations that are sufficient to get across the idea to be learned, but from which distractions and complexities have been reduced. Finally, because it assumes that individual trainees are fully qualified individually on their weapon systems, DIS is often inappropriate for use in training basic job skills regardless of the level of fidelity.

Papers suggested for further reading on topics related to Question 5 include Hays and Singer (1989, Ch. 3), and Waag (1981).

6. Is it really possible to have too much fidelity?

Yes! Absolutely! As hinted above, too much detail and too much realism can get in the way of learning basic skills and procedures, particularly in the earliest first stages of training on a given task. Real world tasks tend to be complex, and in order to help the trainee understand how the task procedures basically work, it is often advantageous to strip away some of the complexity, simplify the task, and even to stop and replay task situations (or use similar techniques) to convey to the trainee what he is supposed to do and how he is supposed to do it. However, although a relatively high level of fidelity might be a handicap during certain early stages of learning, it might also be a near-absolute requirement during later stages.

Thus, in designing a training system, all of the resources that could be used in the desired training should be taken into account, with each of the resources considered for use as part of a complete or full training system. Then, the final design should be based on selections of resources that tend to optimize (if not maximize) training cost-effectiveness if used in the ways defined.

Papers suggested for further reading on topics related to Question 6 include Semple, et al. (1981), and Waag (1981).

**7. When I ask about fidelity, people keep telling me, "It all depends."
Why can't I get a straight answer?**

Because it *does* depend--on what you intend to use the simulation for. When you're training things like operating procedures, communication, decision making, and tactics, high fidelity in *vehicular-handling characteristics* doesn't add anything except to the cost. If you're primarily interested in training manual-control skills, the answer is obviously different.

The best way to get a good answer to the fidelity question is to take a closer look at just exactly *what the simulation is intended to accomplish*. A good approach is to build the simulation from the ground up by *matching* the characteristics of each component and subcomponent to a defined purpose or *intended use* of the simulation. Each fidelity decision must be systematically based on a specific *requirement*. We call this approach *fidelity anchoring*. It is less costly and more effective than the older approach, which sought to compute a training-effectiveness value attributable to each fidelity (realism or validity) component, first by constructing the highest level of fidelity (realism or validity) achievable for the entire simulation, and then by omitting or degrading selectively the fidelity components while measuring the training effectiveness obtained.

8. What was that about *fidelity anchoring*?

Fidelity anchoring is an approach to deciding how a simulator or simulation should be designed in order to meet its objectives with minimum cost and complexity. It presumes (not unreasonably) that we can specify *what the simulation is intended to accomplish* and the probable *range of applications* for which it will be used. Given that information, the premise is that each decision about the appearance and operation of the simulation must be justified (or *anchored*) by a systematic examination of requirements. The three *anchoring criteria* that can justify increases or decreases in fidelity or realism of a component are *effectiveness*, *user acceptance*, and *affordability*. *Effectiveness*, much like *validity*, has to do with how well the simulation does what it purports to do--that is, how well it achieves its intended purpose.

The three criteria are related to one another, but not in any simple way. For example, raising the fidelity of a component is almost always associated with *higher cost*, and, therefore, *lower affordability*, but might not change *effectiveness* or *user acceptance* at all. In general, increases in fidelity will be associated with increases in user acceptance

only where the available fidelity is marginal for the specific application in question. Effectiveness may be associated with user acceptance, but only over a very narrow range.

The objective of fidelity anchoring is to ensure that every component of a simulation--every aspect of how it looks, feels, and operates--should have the exact degree of fidelity required by its intended application, but no more and no less. To design properly the simulators and other training devices to be included in a total training system, one should take into account not only the aspects of the training system, but also those of the operational plan for weapon system usage. For example, if the operational use of an aircraft system specifies that pilots are to fly so many missions per month, with so many takeoffs and landings, it may be that such operational use will provide sufficient practice in executing takeoffs and landings to preclude the need for additional practice of that function in simulators or other training devices within the total training system.

9. OK, I understand now what fidelity anchoring is intended to do. How does one go about doing that?

The process of fidelity anchoring is outlined briefly over the next few pages. While the process appears to be complicated, it is, in reality, only a sequence of relatively straightforward decisions. It is by no means necessary to follow the full development of fidelity anchoring to appreciate its logic. The essential idea is that *decisions* on both *configuring* a simulation system and *investing* the resources for it should be based on a systematic rational examination of *how that specific simulation is to be used*. In that context, generalities about the pros and cons of high and low fidelity are not very helpful. Instead, fidelity anchoring calls for a detailed examination of the simulation requirements on four key dimensions or *fidelity drivers*. *Simulation requirements* are then systematically analyzed and cross-compared to derive *fidelity requirements*. The four key fidelity drivers are identified and discussed below and on the following pages:

- a. **Mission(s) or Mission Segments To Be Simulated.** For realistic practice or evaluation to occur, the system must be used to perform some *mission*. We may wish to simulate all mission phases or (frequently) only selected *segments* of a mission; for example, it may be unnecessary to simulate in full detail routine actions during extended periods of transit time to a battle area.

For example, Figure 1 shows a night interdiction mission profile for the A-6E aircraft, with 16 mission segments. For a given simulation system, we may

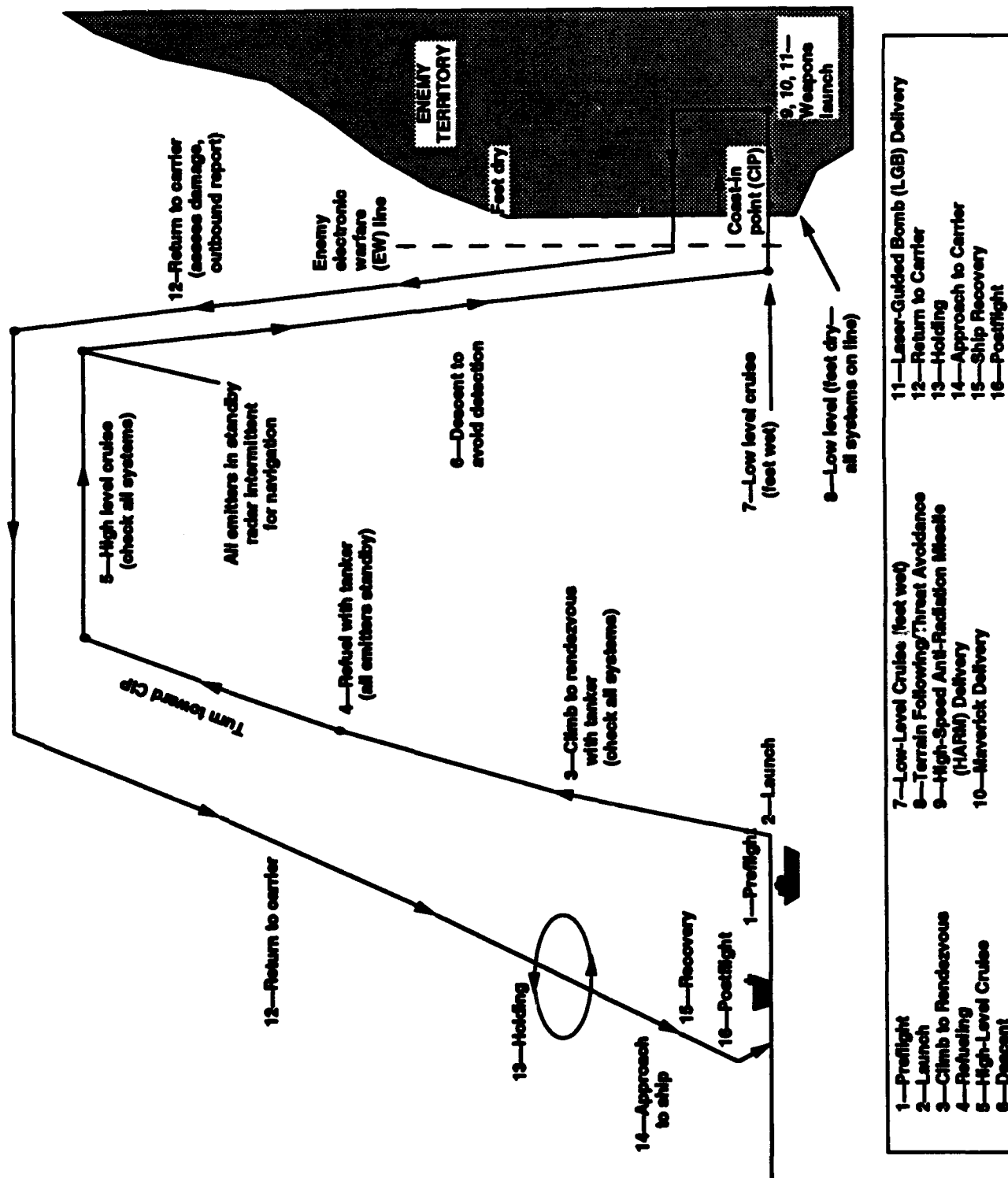


Figure 1.—A Representative Mission Profile

need the capability to simulate (a) all of these in detail, (b) only those that involve potential interaction with specific classes of threats, or (c) any other set of interest.

The *segments* that we elect to include in the simulation will dictate the specific *tasks* to be performed by the operator, the *system components* involved in performing these tasks, and thus the *simulation components* on which fidelity anchoring should be focused.

- b. **Objectives of the Simulation.** A simulation is intended (a) to provide practice on specific skills, (b) to reinforce acquisition and use of job-relevant knowledge, or (c) to evaluate a system or a new weapons concept. These potential *objectives* can be described in terms of broad classes of basic *operator activities* that the simulation does or does not need to *support* (i.e., provide a capability to perform).

Figure 2 lists representative activities or objectives, some or all of which (including others from different lists) might be required to perform tasks during selected mission segments. There are two distinct kinds of activities: (a) *platform-related activities* are those that involve the operation, control, and maneuvering of a vehicle or single weapon system, and (b) *mission-related activities* are those required to employ that weapon system in a *battlefield context* as part of a *coordinated effort* with other platforms in response to a common threat--in short, activities involved in practice or demonstration of *warfighting skills*.

The fidelity required to meet specific objectives would be based on the extent to which each of the identified activities that occur within a mission segment must be supported by the simulation, and in what detail. Note that activities must be cross-referenced to segments. Not all activities will occur in all segments, and the fidelity required by a given activity may differ across segments.

We have thus far considered two of the four fidelity drivers--two factors of fidelity anchoring: (a) the *mission(s) or mission segments* to be performed, and (b) the *objectives* to be achieved, defined in terms of the *operator activities to be exercised*. The remaining two of the four are identified and discussed as follows:

- c. **Fidelity Dimensions.** The dimensions on which fidelity can be examined and evaluated can be grouped into three general classes: there are dimensions that show and describe the attributes of (a) the simulator, including its workstation and its task environment, (b) the operator or team tasking, and (c) the processes or events external to the simulator itself.

Figure 2. Simulation Objectives (Activities To Be Trained, Practiced, or Evaluated)

Platform-Related Activities

- Using Procedures and Principles
- Using Discrete Motor Skills
- Using Continuous Motor Skills (Psychomotor)
- Making Perceptual or Sensory Discriminations
- Communication (internal)

Mission-Related Activities

- Familiarization
 - Terrain
 - Threat Characteristics or Behavior
 - Applying Tactics or Rules of Engagement
 - Planning
 - Mission Profiles
 - Logistics (Fuel, Ordnance, etc.)
 - Coordination with Other Platforms
 - Communication (Internal and External)
-

Figure 3 lists some of the characteristics that make up these three classes of fidelity dimensions. Although all three classes must be considered in fidelity anchoring, the distinctions among the classes are particularly critical for DIS.

The first class of attributes describes the (simulated) system itself and how it functions as a freestanding entity independent of any simulation network; the concerns for fidelity here are those that address the *operator's equipments and its immediate environment*—the *look* and *feel* that are the results of the physical, sensory, and perceptual variables employed. These are basic aspects that need to be considered for all simulations, conventional or DIS.

Tasking attributes are drivers that determine the specific tasks to be performed by the operator and the task loading under which he will work. Note that these drivers are external to the system--what the operator does in the simulation is determined by how the simulated system is used. These drivers are present in both DIS and conventional simulation. In conventional simulation they are likely to be built into the simulator, whereas in DIS they are more likely to be heavily influenced by the activities of other entities.

Figure 3. Fidelity Dimensions

Some Attributes of the --

-- Simulator

- Appearance of equipment ("looks like")
- Discrete operation of system (switches, discrete controls)
- Continuous operation of system (dynamics, control response, feedback)
- Sensory stimulation (motion, visual, sound)
- Physical environment (heat, cold, vibration, illumination)
- Perception of simulation by human (discriminability of stimuli and cues)

-- Tasking

- Task characteristics (cues, procedures)
- System employment (force organization, logistics)
- Demands of the operator (loading, pacing, OPTEMPO)

-- Simulation Process

- Characteristics of other operators (experience, training, aptitude)
 - Characteristics of threats (emissions, signatures, capabilities, numbers)
 - Behavior of threats (deployment, tactics, rules of engagement)
 - Characteristics of own forces (emissions, signatures, capabilities, numbers)
 - Behavior of own forces (tactics, rules of engagement)
-

The attributes of external processes generally arise from the dynamics of system participation in the interactive events of the battlefield. Although some of these attributes are present in conventional simulation, they are much more complex and much richer in DIS. Each of these attributes and characteristics defines a dimension of fidelity, and each requires decisions about the level of fidelity at which that attribute will be represented for a component of the simulation within the context defined by the missions and objectives of the desired training.

- d. **Simulation Components.** As suggested by the prior discussion, a simulation can be thought of as one or more simulators interacting as required by the evolving scenario. In the larger context of DIS, the process may involve not only simulators, but also actual equipment and computer models (automated forces and evaluation rules).

Execution of a simulation requires the presence of numerous simulation components. Figure 4 shows one way of breaking down a DIS simulation process into a set of components. Note that these are logical, but not necessarily physical components; that is to say, there may not be a specific box or piece of equipment that corresponds to each identified component.

Note also that the components are subdivided into local and global sets. Local components are part of the simulator and its immediate environment; they have roughly the characteristics defined as *attributes of the simulator* in Figure 3. Global components are defined by the external processes and environments with which the local components interact. The importance of breaking down a simulation into its building blocks is that it is about these individual components that fidelity decisions must be made. Thus, fidelity anchoring is applied to decide *which components require what kind and level of fidelity*.

10. So we have four sets of dimensions. How do we use them to make fidelity decisions?

The four *key fidelity drivers* just discussed constitute four dimensions on which any given simulation can be analyzed. In the analytic process of fidelity anchoring, these four dimensions are combined with three criteria to lead us to appropriate simulator-fidelity decisions. Two criteria, *training effectiveness* and *user acceptance*, are employed in the earlier stages, whereas the third criterion, *affordability constraints*, becomes relevant only after we have identified alternative possible simulator configurations. Thus, we complete the process of fidelity anchoring by cross-comparing the alternatives on certain combinations of these dimensions and criteria. Our ultimate objective is to determine, for each *component* (as in Figure 4), on each *fidelity dimension* (as in Figure 3), the *degree of fidelity* required to support the intended uses of the simulation.

The *process of fidelity anchoring* involves four stages as follows:

- a. In the first stage, we determine which fidelity dimensions are relevant to each simulation component. Consider a two-way matrix formed by the intersections of the 24 simulation components (see Figure 4) with the 14 fidelity dimensions (see Figure 3). The result is a matrix of 336 cells that could be examined. For each cell, we must make a relevance judgment. While this seems initially to be a large number of relevance decisions, in most cases only a very limited set of cells will actually be examined, because not all fidelity dimensions are associated with all components. For example, in evaluating the workstation component, we are not concerned with any attributes of the simulation process or with most of the

Figure 4. Components of a Distributed Interactive Simulation (DIS)

LOCAL — Simulator

- Workstation
 - Controls/Displays
 - Models or Equations of Vehicles and Weapons
- Physical Environment (Atmosphere, Cold, Heat, Light, Vibration, etc.)
- Sensory Environment (Auditory, Visual, Motor, etc.)
- **Database (Local)**
 - Gaming Area
 - Terrain
 - Object Representation (State, Location, Attributes)
- **Operator(s)**
 - Aptitude
 - Training
 - Experience
- **Mission Characteristics**
 - Procedures
 - Tactics/Rules of Engagement
 - Profiles/Scenarios

GLOBAL — Other Entities

- Actual Equipment
 - Threats
 - Own Forces
 - Simulated
 - Threats
 - Own Forces
 - Semi-Automated/Automated
 - Threats
 - Own Forces
 - **Database (Global)**
 - Object Representation (State, Location, Attributes)
 - **Network**
 - Protocols
 - Object and Entity Data
 - State Changes
 - Location Changes
-

tasking attributes, and thus need to make relevance judgments over no more than a dozen or so cells. Similarly, relevance judgments will be required over only a small fraction of the total 336 intersection cells. Only the relevant cells will be considered in later sequences of fidelity anchoring.

- b. In the second stage, for each relevant combination identified in stage 1, we determine the highest fidelity required to attain the desired simulation objectives (see Figure 2) in any mission segment to be simulated (see Figure 1) on either of two criteria (effectiveness and user acceptance). In other words, for a given fidelity dimension-component combination, we look through all the possible things that we might want to train or evaluate, and we look across all segments of all missions to be trained, and decide on (make a judgment regarding) the highest level of fidelity that might be needed in the intended use of the simulation. Further, we make this judgment on either of two bases--a given level of fidelity may be required either to make the simulation sufficiently effective or to achieve a desired level of user acceptance independent of effectiveness--and we take the higher of the two. The fidelity required could be assessed on a variety of scales, but the simplest scale would likely be the three-point rating, *High, Medium, or Low*. After all this work, we end up with a single data point for each relevant combination, the highest fidelity that we found in any simulation objective evaluated across segments and across criteria. We repeat the operations of Stage 2 for every relevant combination from Stage 1, and for each of the missions that might be of interest, and then move on to Stage 3 with a collection of High, Medium, or Low judgments for these combinations.
- c. In the third stage, we return to our Stage-1 two-way matrix (composed of fidelity dimensions crossed with simulation components) and enter into each relevant cell the highest fidelity judgment we determined in Stage 2, and use those entries in deciding how to go about designing each component to have exactly the appropriate level of fidelity on each fidelity dimension. For example, given our look at objectives, missions, and criteria, we might have determined that the simulation's workstation component requires a high level of fidelity on discrete operation of equipment, but only medium fidelity on appearance of equipment, and a low level of fidelity on threat characteristics. We may find that there are no requirements for manipulation of the simulator's physical environment, and that the sensory environment requires medium fidelity visual, but with no requirement for motion or sound. Thus, by going through the first three stages, we have anchored our fidelity judgments in a orderly, systematic examination of what the simulation is intended to do.

d. In the fourth and final stage, we introduce our third criterion for fidelity of a simulation--affordability. After we have decided through careful analysis what our simulation characteristics should be, we may find that we cannot afford some of the desired levels of fidelity. Indeed, some of the intended uses of our simulation may require us to network numerous simulators, and it may be necessary to compromise some of our fidelity requirements in order to afford the additional simulator-units--that is, to reserve the resources to acquire simulator units in the required quantities. Again fidelity anchoring is useful. By the end of Stage 3, we know rather precisely those components, missions, and objectives that are the major fidelity (and thus cost) drivers, and can work backward from our Stage 3 matrix to identify where requirements may be eased with least loss of usefulness for the simulation and in achieving its objectives.

11. I'm confused about the logic of using the *highest* fidelity requirement. Aren't we trying to *reduce* unnecessary fidelity?

It's true that the reason for going through a process like fidelity anchoring is to make sure that a component has the kinds and degrees of fidelity appropriate to that for which we want to use the simulation. However, if we examine what goes on in Stage 2, we will see that the logic of identifying the highest fidelity requirement is carried out on combinations of objectives and mission segments. Once we have determined a simulation configuration, that is the configuration that must be used for *all* the included objectives and mission segments to be addressed; that is, the fidelity of a component cannot be changed as the simulated mission moves from one segment to another. Even if only one objective or mission segment in our intended use requires a given high level of fidelity, that is the level that would have to be provided--even if every other objective and segment called for only low levels of fidelity. Fortunately, the process of fidelity anchoring will identify such imbalances where they occur, and we can then reevaluate the situation to decide whether we really need to include the objective or mission segment that requires the high-level fidelity.

12. We've spent a lot of time on fidelity anchoring. Why are you so concerned about it?

In general, the higher the fidelity needed for a simulation, the more it costs to design, build, and maintain. In other words, the higher the fidelity of a simulation, the higher its life cycle costs. Increases in life cycle costs are *not* proportional to increases in fidelity. Small increments in fidelity can multiply costs dramatically. In addition, the increased system complexity associated with achieving higher fidelity is associated with

decreases in the reliability and the availability of simulation equipment. Given these conditions, we obviously do not want to pay for any more fidelity than we need. Fidelity anchoring is one way of forcing a systematic justification of fidelity requirements and ensuring that fidelity decisions are based on identified uses of the simulation.

Not having to pay for more fidelity than you need, clearly important for conventional simulations, is even more critical for DIS applications. DIS is most valuable for large-scale, multiple-entity simulations suitable for the training of warfighting skills through coordinated unit (collective) training, or for evaluating how proposed (or actual) weapon systems or tactics perform in the context of a realistically simulated battlefield. Realistic two-sided free-play exercises using DIS could involve dozens, and perhaps hundreds or thousands of workstations. Other than actual field exercises, DIS is the only way to accomplish some of these very large-scale training or evaluation objectives. To make such DIS exercises possible, simulations and their components must be affordable--and in sufficient quantities to permit achievement of the objectives of their use. Systematic attention to fidelity is key to resolving affordability issues.

13. I notice that fidelity anchoring uses relative judgments of fidelity requirements (high, medium, or low). Isn't there some better way to quantify fidelity?

At present, fidelity is a metric-free construct with no agreed-upon measurement scale on which the fidelity of a specific simulation can be located and assigned a numerical value. Fidelity is not a unitary concept; it has many dimensions and subdivisions--far too many to be represented by any currently recognized single index. The problem is addressed in fidelity anchoring by examining the fidelity requirements, not of the simulation as a whole, but of specific simulation contexts; for example, a specific context defined by one mission segment and one objective. In such contexts, only one dimension at a time is of concern (e.g., how much fidelity is necessary for a specific use) and for that purpose it is possible to make meaningful relative judgments (e.g., *medium* represents less fidelity than *high*, but more than *low*). We might even be able to convert those judgments into numerical indices, but they would have meaning only in that defined, limited context, not for the simulation as a whole.

There are two problems with having to make relative judgments (high-medium-low) about fidelity requirements. First, they are subjective, and different people might come up with different values. There is not much we can do about that. It is really not much

different from the educated guesses that we must frequently make in simulator design today. However, the extra advantage of fidelity anchoring is that it forces us to take some major drivers into explicit account. It also breaks down the simulation into simple, specific contexts that make the judgments more accurate. Secondly, relative judgments do not really tell us exactly how the specific components and subcomponents should look and operate. They only tell us that a component needs a lot of fidelity (or a little or something in between). If necessary, we could expand our relative judgment scale with verbal anchors that say more precisely what we mean by the terms (e.g., *low* means a non-functional mockup can be used, while *high* means that the use of actual equipment is required). We could even have several different scales, one dealing with appearance, one with operation, etc. The purpose of fidelity anchoring is to focus our attention on the things that matter; it will not make automatic decisions for us. The technology for that, simply, is not yet available.

We are currently working on some new ways of better quantifying fidelity requirements for some very specific objectives and components. For example, in simulating vehicular control, we want to provide just enough fidelity (through equations of motion, information content, and update rates) for the operator to be able to employ the same procedures, cues, and motor responses used in the actual system. When we reduce fidelity below some low point at which the operator must deviate significantly from established control strategies, we have essentially made the control of the simulated system a different task from control of the actual system, and run the risk that some of our objectives might not be met as well or even at all. If we could define that critical low point quantitatively, we would have a way of evaluating a given simulation with respect to its fidelity for vehicular control. Although the way we quantify a simulation for vehicular control does not necessarily generalize to other objectives, it does illustrate an important point about fidelity determination; namely, that we want to be able to identify precisely that level of fidelity which *minimizes the need for task redefinition* throughout the simulation--that is, the level that allows operators to perform the tasks required by the simulation in much the same way, using the same information, and with the same speed as the equivalent tasks performed in the actual system.

For further reading on topics related to Question 13, see Meister (1990).

14. That's enough about fidelity for now. Let's backtrack a bit. You keep making a distinction between basic skills and warfighting skills. What's the difference and why does it matter?

Basic skills are all the things the operator has to learn to do his or her job--that is, to become qualified to operate the system or platform. *Basic* does not necessarily mean simple. Some of these "basic" skills are in fact very complex; they include things like flying an aircraft or operating a radar tracking system. But they are primarily *platform proficiencies*--that is, concerned with a single platform or system, or perhaps a few systems operating together, and are largely mastered by the end of regular training. On the other hand, *warfighting skills* have to do with the use of the system in a collective military operation--for example, as part of a force-level or even a theater-level response, in coordination with other air, ground, and sea forces. There are required skills (tasks, tactics, and procedures) that go beyond the operation of an individual system or unit in this larger battlefield environment. The players in this simulated battlefield environment are not only the weapon system operators, but also the commanders, staffs, logisticians, support units, intelligence personnel, and decision makers at all levels--in short, all the combat, combat support, and combat-service support elements assigned to the battle force and its support. By definition, such warfighting skills are acquired and practiced in large-scale exercises and simulated or actual war. The large-scale exercises require actual maneuvers with a great many people and systems; the simulated battlefields require the kinds of distributed interactive simulations addressed here.

The training applications of *Distributed Interactive Simulation* (DIS), as we have used the term, are intended nearly exclusively for the training and practice of warfighting skills. DIS is not designed for, and may not always be appropriate for, basic skills training. For example, it assumes that a pilot knows how to fly--how to control and operate his aircraft--and is not intended to develop or improve the pilot's basic airwork or flying skills. Accordingly, fidelity concerns for an aircraft simulated in a DIS configuration could be quite different from those of a simulator intended to provide airwork training. Similarly, as we know from the SIMNET program, skills such as planning, coordination, and communication in armor operations can be trained effectively in a simulator that does not look or operate exactly like a real tank. It is important to remember that the *intended use of the simulation* is the only acceptable basis for decisions about fidelity, validity, realism, and all the other terms that apply to a simulation's correspondence with reality. The intended uses of DIS are almost always different from those of simulators used in

basic skills training, and these distinctions must be kept clearly in mind when evaluating a simulation for either use--DIS or basic skills training.

Papers suggested for further reading on topics related to Question 1 include Alluisi (1991), Thorpe (1987), and Madden (1988).

15. If fidelity requirements are only dependent on intended use, why do people worry so much about *negative training*?

When people talk about *negative training*, they are probably referring to what skill acquisition theory calls *negative transfer*. Negative transfer from simulation would mean that something about a task learned in a simulator would actually interfere with the ability to perform that task in the real world; that is, the task would be performed after the simulator practice less well than without such practice. Negative transfer in today's simulation world is almost nonexistent. Simulators typically provide what we call positive or partial transfer; some portion of the simulated practice goes to improve performance on the real task. This is sometimes measured in transfer effectiveness ratios (TERs). A TER of 0.8 would mean that 5 hours in a simulator gives the same improvement on the job as 4 hours on the actual system; cases where the TER is greater than unity are also known.

TERs are sometimes evaluated on the basis of how much it costs for one hour in the simulator versus one hour in the real-world operating system. A simulator with a TER of 0.5 is a real bargain if the system costs 3 or 4 times as much per hour to operate as the simulator. In modern simulation, we would be concerned if a TER was much lower than 0.3 or 0.4. Transfer is almost never zero, and negative values are extraordinarily rare. To get negative, as contrasted with partial, transfer, the simulator would have to teach something that dramatically interfered with performance on the real-world actual system--for example, controls operating in reverse, incompatible procedural steps, etc. *Reduced or lower fidelity* refers only to the completeness with which a system is represented. If the time needed for training in a given simulator or training device is greater than in the actual equipment, that training equipment is by definition of reduced or lower fidelity. This does not imply an incorrect representation, a condition that should never occur in a properly designed simulation. Reduced fidelity simulations do not cause negative transfer.

Papers suggested for further reading on topics related to Question 15 include Cormier and Hagan (1987), Hays and Singer (1979), Rankin et al. (1984), and Waag (1981).

16. You mentioned earlier the *validity* of a simulation as another way of talking about fidelity. Can you expand on that idea?

Previously, we distinguished among the terms fidelity, validity, and realism. Although, for convenience, we deliberately ignored some of these distinctions in some answers, the differences are important for understanding current thinking about fidelity issues. We have made the point that fidelity as a term has no useful meaning until we ask, "Fidelity for what purpose?" Once we have defined the intended use, we are ready to talk about the *validity* of a simulation. In our earlier discussion, we suggested that the term fidelity might usefully be reserved for descriptions of the degree of physical correspondence between a simulation and the reality it represents, while using the term validity to refer to the extent to which the simulation achieves its purposes. We noted that there were at least three kinds of applications for simulations--training, acquisition, and analysis. Each of these would have different goals, and the validity of any one simulation used in the different applications would vary accordingly. A simulator with high or low validity for one use would not necessarily have the same degree of validity for another use. Focusing on the term validity is one way of emphasizing a point that we made previously; namely, that fidelity is not an objective or a driver for a simulation, but rather it is only a way of configuring the simulation in order to achieve a desired validity.

Validity as a term shares a common difficulty with fidelity--neither is easily or readily quantifiable. In its current usage, validation is seen not as an event or result with an associated number, but rather as an ongoing process or set of operations that evaluates the correspondence between simulation and real world from the perspective of intended use. That is to say, a simulation is never *validated*, but rather it undergoes *validation*. Thus, the validation process involves the gathering of evidence about credibility. This evidence can be in many forms, from simply verifying that the logic and assumptions used in the simulation design are "reasonable," to evaluations of simulation outcomes by subject matter experts, to formal comparisons of outcomes with laboratory findings or operational data. The more we can obtain such evidence about a simulation, the more credible our results and the higher its judged validity. Because simulations have assumed such an important role in all aspects of skill training, analysis, and system design, test and evaluation, there has been considerable interest in recent years in standardizing the definitions of terms used to describe validation operations, and in distinguishing them from terms that describe other kinds of simulation evaluations. Two such terms are *verification* and *accreditation*. *Verification* of a simulation is simply the process of determining that the simulation is

properly implemented; that is, that it meets design specifications and that its equations and algorithms are correctly translated computer codes. *Accreditation* is a formal decision, given the weight of evidence, that the simulation is sufficiently valid to be acceptable for a specific intended use. Procedures for formal accreditation vary by Service and are usually implemented by a formally published instruction.

Papers suggested for further reading on topics related to Question 16 include Hodges and Dewar (1992), Jones et al. (1985), and Williams and Sikora (1991).

17. OK, that just about wraps it up. Any last comments?

Just a few. Over the past 30 years, simulation has become both an increasingly viable and a progressively more necessary alternative to using actual equipment and field exercises for training personnel in warfighting skills and for carrying out certain phases of weapon system design and evaluation. These trends are highly likely to continue and probably accelerate, given the steady improvements in all aspects of simulation technology. A well-designed simulation can certainly offer dramatic leverage on cost-effectiveness compared to most other ways of achieving the same objectives. As projected drawdowns in equipment and personnel occur, simulation will eventually represent a most attractive (if not the only feasible) mechanism for certain kinds of training and system evaluations, and the technology of distributed interactive simulation will represent a most attractive (if not the only practical) avenue by which such simulations can be carried out.

When operations that have been anchored in the real world are replaced by those removed a step further from reality, there is always a concern that something crucial has been lost. It is important to understand that what is lost in moving to simulation can be minimized by careful attention in simulation design to replicating in greatest detail those features that are most significant for the intended application. For this sort of tailored design to take place, we must be able to see fidelity and validity, not as indivisible, all-or-none concepts, but rather as context-specific terms with many levels of gradation. Generalities are not useful. We must be thoughtful and precise in our evaluation of a simulation, always forming our questions in the context of intended use. Over time, it is inevitable that simulations will encroach on a host of new application territories. The decisions to be made will probably relate less to whether or not simulation should be used, but rather more with how simulation can be most successfully implemented.

REFERENCES

- Advisory Group for Aerospace Research and Development [AGARD] (1980). *Fidelity of simulation for pilot training*. AGARD AR-159. Neuilly-sur-Seine: Author.
- Alluisi, E.A. (1991). The development of technology for collective training: SIMNET, a case history. *Human Factors*, 33, 343-362.
- Blaiwes, A.S., and Regan, J.J. (1986). Training devices: Concepts and progress. In J.A. Ellis (Ed.), *Military contributions to instructional technology* (pp. 83-170). New York, NY: Praeger.
- Cormier, S.M., and Hagman, J.D. (Eds.). (1987). *Transfer of learning: Contemporary research and applications*. Orlando, FL: Academic Press.
- Cream, B.W., Eggemeier, F.T., and Klein, G.A. (1978). A strategy for the development of training devices. *Human Factors*, 20, 145-158.
- Hays, R.T. (1981). A fidelity definition and outline of interactive factors. In R.T. Hays (Ed.), *Research issues in the determination of simulator fidelity: Proceedings of the ARI-sponsored workshop* (TR-547; pp. 65-76). Alexandria, VA: Army Research Institute.
- Hays, R.T., and Singer, M.J. (1989). *Simulator fidelity in training system design: Bridging the gap between reality and training*. New York, NY: Springer-Verlag.
- Hodges, J.S., and Dewar, J.A. (1992). *Is it you or your model talking? A framework for model validation* (R-414-AF/A/OSD). Santa Monica, CA: The Rand Corp.
- Institute for Simulation and Training [IST]. (1990). *Summary Report. The third workshop on standards for the interoperability of defense simulations. Vol. I: Minutes from plenary session* (IST-CR-90-13). Orlando, FL: University of Central Florida.
- Institute for Simulation and Training [IST]. (1991). *Military Standard (Draft). Entity information and entity interaction in a distributed interactive simulation* (IST-PD-90-2) Orlando, FL: University of Central Florida.
- Jones, E.R., Hennessy, R.T., and Deutsch, S. (Eds.). (1985). *Human factors aspects of simulation*. (Report of the National Academy Working Group on Simulation). Washington, DC: National Academy Press.
- Kinkade, R.G., and Wheaton, G.R. (1972). Training system design. In H.P. Van Cott and R.G. Kinkade (Eds.), *Human engineering guide to equipment design*. Washington, DC: Government Printing Office.

- Madden, J. L. (1988). Global strategy for the simulation network. In A. C. F. Gilbert (Ed.), *Proceedings of the 30th annual conference of the Military Testing Association*, Arlington, VA, December, 1988 (pp. 849-854).
- Meister, D. (1990). *Conceptual aspects of human factors* (pp. 145-163). Baltimore, MD: The Johns Hopkins University Press.
- Miller, R.B. (1954). *Psychological considerations in the design of training equipment*. (Tech. Rep. No. 54-563). Wright-Patterson AFB, OH: Wright Aeronautical Development Center.
- Orlansky, J., and String, J. (1979). Cost effectiveness of flight simulators for military training. *Proceedings of the Interservice/Industry Training Equipment Conference*, Orlando, FL (pp. 9-20).
- Rankin, W.L., Bolton, P.A., Shikiar, R., and Saari, L.M. (1984). *Simulator fidelity and training effectiveness: A comprehensive bibliography with selected annotations* (NUREG/CR-3726). Washington, DC: Nuclear Regulatory Commission.
- Ropelewski, R.R. (1990). Army takes SIMNET for training: DARPA seeks new users, new uses. *Armed Forces Journal International*, March 1990, 66-68.
- Semple, C.A., Hennessy, R.T., Sanders, M.S., Cross, B.K., Beith, B.H., and McCauley, M.E. (1981). *Aircrew training devices: Fidelity features* (AFHRL-TR-80-36). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.
- Su, Y.-L.D. (1984). *A review of the literature on training simulators: Transfer of training and simulator fidelity* (Rep. No. 84-1). Atlanta, GA: Center for Man-Machine Systems Research, Georgia Institute of Technology.
- Thorpe, J.A. (1987). The new technology of large scale simulator networking: Implications for mastering the art of warfighting. *Proceedings of the Ninth Interservice/Industry Training Systems Conference* (pp. 492-501). Washington, DC: American Defense Preparedness Association.
- Waag, W.L. (1981). *Training effectiveness of visual and motion simulation* (AFHR -TR-79-72). Williams Air Force Base, AZ: Air Force Human Resources Laboratory.
- Williams, M.L., and Sikora, J. (1991). SIMVAL minisymposium--a report. *Phalanx*, 24(2), 1-6.