RAND

Enhancing Weapon System Analysis

Issues and Procedures for Integrating a Research and Development Simulator with a Distributed Simulation Network

Monti Callero, Clairice T. Veit, Eugene C. Gritton, Randall Steeb

Prepared for the Advanced Research Projects Agency

National Defense Research Institute

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank,	2. REPORT DATE	
	1994	TECHNICAL REPORT

4. TITLE AND SUBTITLE
ENHANCING WEAPON SYSTEM ANALYSIS ISSUES AND PROCEDURES FOR INTEGRATING A RESEARCH AND DEVELOPMENT SIMULATOR WITH A DISTRIBUTED SIMULATION NETWO

5. FUNDING NUMBERS

Callero Monti, Veit Clairice T.,Gritton Eugene C.,
Steeb Randall

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
RAND
NATIONAL DEFENSE RESEARCH INSTITUTE

8. PERFORMING ORGANIZATION REPORT NUMBER
DSRS U 35253

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
ADVANCE RESEARCH PROJECT AGENCY
3701 FAIRFAX DRIVE
ARLINGTON, VA 2203

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
MDA90390C0004

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
THIS REPORT ADDRESSES THE FEASIBILITY OF APPLYING SIMULATOR NETWORK TECHNOLOGY TO WEAPON SYSTEM ANALYSIS IN SUPPORT OF THE WEAPON SYSTEM DEVELOPMENT AND ACQUISITION PROCESS. THE REPORT SHOULD BE OF INTEREST TO PERSONNEL IN THE MILITARY SERVICES AND DEFENSE OFFICES AND AGENCIES INVOLVED IN THE WEAPON SYSTEM DEVELOPMENT AND ACQUISITION PROCESS AND IN MODELING AND SIMULATION ACTIVITIES.

14. SUBJECT TERMS
weapon
simulation

15. NUMBER OF PAGES
22

16. PRICE CODE

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

Preface

This is the final report of a project conducted for the Advanced Research Projects Agency that addressed the feasibility of applying simulator network technology to weapon system analysis in support of the weapon system development and acquisition process. The research was carried out in the Applied Science and Technology Program within RAND's National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense and the Joint Staff.

The report should be of interest to personnel in the military services and defense offices and agencies involved in the weapon system development and acquisition process and in modeling and simulation activities.

Contents

Appendix

Figures

Tables

Summary

In this report we propose that operational effectiveness analyses supporting the weapon system development and acquisition process would be enhanced by properly integrating highly realistic, human-operated research and development simulators of the emerging system with distributed interactive simulation (DIS) networks that provide large-scale, human-in-the-loop battlefield environments. We further propose and describe how to approach an initial integration that would both serve as a proof-of-principle experiment and provide intellectual, hardware, and software "tool kits" for future integrations. Our approach emphasizes aspects of simulator integration such as fidelity, validation, correlation, and calibration that are critical to weapon system analyses but are not being undertaken in current efforts to create DIS environments simply through communication interfaces of disparate simulators with SIMNET-based systems.

For this proof-of-principle demonstration, we recommend integration of the Crew Station Research and Development Facility (CSDRF) rotorcraft simulator operated by the Army and NASA with a fidelity-enhanced version of the Army Aviation Test Bed (AVTB) that incorporates the SIMNET-based simulator networking technology developed by the Advanced Research Projects Agency (ARPA). We discuss what such an integration could contribute, describe what we believe to be a feasible approach for integrating the two systems, and discuss the important technical and validation issues associated with such an integration, including the requirement to improve the fidelity of many existing AVTB/SIMNET features.

Our main interest in this research stems from the proposition that the weapon system development and acquisition process should have solid analytical support from beginning to end so that, to the degree possible, design decisions to maximize operational effectiveness with minimum cost are made well before the item is actually built. The analytical methods currently available lack important features for assessing the capabilities of future weapon systems in battlefield operations. Three primary analytical methods are now used to evaluate new weapon systems—computer simulations, research and development simulators, and live operational tests and evaluations. Each of them has important constraints in its ability to properly represent weapon systems' operations in expected combat environments. For example, representations of human

behavior (e.g., decisionmaking, operator actions) are notably poor in computer simulations of combat, the primary tool used in most weapon system analyses; human-operated research and development (R&D) simulators have limited representations of combat environments; and live operational tests and evaluations feature constrained operations in limited combat environments, at high costs. Better analytical tools and methods are needed to assess new weapon systems in large-scale combat operations that allow for human-in-the-loop participation in developing combat tactics and decisionmaking. We believe that such capabilities would lead to an enhanced capability for weapon system effectiveness analyses supporting the development and acquisition process.

Enhancement Through Simulator Networking

An important advancement relevant to these needs has been made in simulator networking. ARPA and the Army developed a simulator network featuring distributed processing techniques that allow human-operated simulators and automated (or semi-automated) weapon system simulations to interact on a single simulated battlefield. The system, called SIMNET,[1] has several unique features that would enhance weapon system analyses. Each SIMNET battlefield element (e.g., tank, aircraft, command and control node) is computationally self-sufficient; hence, there is no technical limit on the number and type of interacting human-operated weapon system simulators and automated systems. This, along with semi-automated, unmanned weapon system simulations, allows for large-scale battlefield environments to be created in which to perform analysis. SIMNET includes an extensive set of analytic tools, including a "stealth" vehicle that can observe the battle from various points (e.g., looking out the windscreen of the cockpit of an aircraft) and that has a complete playback capability for all or part of the simulation.

Simulator networking appears to have a large potential for providing a basis to resolve many of the critical deficiencies found in current weapon system analysis methods. However, we have found that simulator networking as currently represented by existing SIMNET systems will require considerable improvements to achieve the required level of capability, particularly in the area

[1]SIMNET and SIMNET derivatives remain the only functional implementations of simulation networking and provide the foundation for the extensive concept and technological development activities in distributed interactive simulation currently under way within the defense modeling and simulation community. For simplicity, we use the term SIMNET without deprecation to other named versions.

of weapon system simulation fidelity.[2] Meanwhile, defense industries and the military service laboratories now routinely develop high-fidelity R&D simulators to test design concepts, investigate human/system interfaces, and conduct other research and development activities on specific emerging weapon systems. The concept addressed here is to combine the high-fidelity representation of a developing system afforded by its R&D simulator with the combat simulation environment afforded by SIMNET, thereby mutually complementing their individual strengths and weaknesses.

The integrated system should provide an improved capability over current computer simulations to conduct weapon system operational effectiveness analysis. With fully human-operated or mixed human-operated and automated forces on the battlefield and employing real-time command and control, the enhancements could include those shown in Table S.1. These enhancements would make possible both a more complete and a more credible analysis of the weapon system and do it in a more realistic combat environment, featuring real-time interactions among combatants and the ability to evolve tactics and techniques throughout the course of the analysis.

The contributions to operational test and evaluation (OT&E) shown in Table S.2 would follow from the ability to conduct pretest assessments with the improved simulation system to identify the most important areas to investigate, determine

Table S.1

Potential Analysis Enhancements

Environment	Current	SIMNET/R&D
Tactics	Specified, adjusted	Evolved during analysis
Vehicle	Coarse dynamics	Crew interaction
Sensor models	Varying fidelity, notional operability	Varying fidelity, crew operability
Terrain	Contours only	Contours, foliage, and culture
Threat	Variable fidelity, notional operability	Variable fidelity, crew operability
Vulnerability	Engineering estimates	Engineering estimates
Weapon effects	Flyout/statistical	Flyout/statistical
Combat interactions	Stylized	Real-time, human-in-the-loop

[2]The term fidelity refers to logical and physical misrepresentations or "correct" representations of a simulation's modeling process (the simulation's apparent faithfulness to reality), determined through close scrutiny of that process.

Table S.2

Potential Contribution to Operational Test and Evaluation

- OT&E plan assessment
- Early operational assessment
  - — Identification of critical issues
- Pretest investigation
  - — Assess and hone test design
  - — Familiarize participants
  - — Avoid execution flaws
- Compensate for real-world limitations
  - — Safety factors (tactics, weather, night)
  - — Multiple weapon systems employment
  - — Density of threats and targets

deficiencies in T&E designs and implementations, and train personnel to conduct the tests and evaluations. In concept, pretests could be run using a replica of the test environment, including instrumentation and human-operated weapon systems. The simulator network would also permit investigation of performance outside of the safety envelopes required for operating actual systems. And larger and more intricate combat environments could be simulated than could be practically and perhaps feasibly established on the test range. Advanced future capabilities that do not exist yet could also be included in analyses.

These potential direct payoffs to the weapon system development and acquisition analytic process could improve the credibility and usefulness of the analyses and increase the benefits over live tests and evaluations. If such improvements are realized, the potential exists for large savings in weapon system R&D and OT&E costs because of an improved understanding of critical technology drivers, system utility, and experimental test design. Costs for integrated simulation network analyses once a basic DIS environment with sufficient validity was in place would probably range in the millions or low tens of millions of dollars, whereas OT&E tests and changes to a system in late development stages or production typically impose costs in the hundreds of millions of dollars.

Technical Issues

To assess the feasibility of this concept, we chose the Crew Station Research and Development Facility (CSRDF) helicopter simulator operated by the Army and NASA at the Ames Research Center as the R&D simulator to be integrated with the Army Aviation Test Bed (AVTB). We made this choice because of CSRDF's government ownership, representativeness of advanced weapon system simulators, and dedication to combining research with practical application. The CSRDF is a full combat mission rotorcraft simulator whose crew stations have glass cockpits that represent all crew station interfaces for daytime or night adverse weather missions. All displays in the cockpit are reconfigurable, and the pilot has a fiber optic helmet-mounted display that provides a 67° vertical by 127° horizontal field-of-view and an unlimited field-of-regard using the General Electric CompuScene IV image generator. This system exhibits significantly higher fidelity than the AVTB crew stations.

AVTB and CSRDF have different design philosophies and architectures, and their integration presents a number of difficult technical problems, all of which must be resolved to provide the needed analytic framework for the weapon system development and acquisition process. In particular, we note that integration to support operations analyses is not achieved simply by intercommunicating between the two systems in the correct protocols; intercommunication is a necessary but not a sufficient element of productive integration. Technical issues to be resolved include

- synchronous and asynchronous processing coordination,
- simulation network communication and protocols,
- database sharing and correlation, and
- sensor modeling and visual image representation.

AVTB processing is distributed and asynchronous. Each simulator has its own processor that is responsible for calculating local vehicle interactions based upon data received from other simulators, and data are transferred in an asynchronous manner among simulators. The CSRDF, on the other hand, is a synchronous system with a single host computer. When the CSRDF interacts with the AVTB environment, it must communicate its status and actions in the same manner as other SIMNET systems, react to incoming messages about another vehicle's status, actions, and appearances, and determine how and when to display the transmitted events.

The integration must use common machine-independent reference databases to ensure that the two systems represent the same terrain characteristics, vegetation, roads, bridges, buildings, and so forth. Database interchange specifications have been developed to provide a basis for interchanging datasets among DIS database users and appear to be sufficient for this task. Both static and dynamic correlation of the databases would be required so that elements appear correctly in both time and position.

AVTB and CSRDF both generate images by synthesizing and displaying scene polygons, but they use very different image generators. Depending on the display being represented (out-the-window, hard optics, TV, forward-looking infrared [FLIR]), there may be large differences between the two simulation systems in the number and type of polygons generated, although both simulators may faithfully reproduce the salient characteristics of the system. The objective is not necessarily to achieve an exact match between reality and simulator-presented image, or between simulators, but to at least produce a sufficient consistency of operator behavior and system performance among the simulators *within* the integrated system and strive for sufficient realism to emulate the performance of the actual systems. In general, all players must appear in the same positions and times in different simulators, and cues for detection, tracking, and other activities must be consistent given clear line-of-sight. Data logging and postprocessing must accurately portray the movements, detections, shots, kills, and status of entities in both AVTB and CSRDF.

System Validation

A major research effort is required to assess the internal and external validity of the AVTB/CSRDF integrated product. Internal validity (or consistency) refers to the performance equivalence among like entities in the integrated system, such as a AVTB helicopter simulator and its AVTB semi-automated counterpart or its CSRDF counterpart. It also refers to process equivalences (or appropriate differences) among *different* systems using the same subsystem (e.g., detection/engagement decisions resulting from an infrared sensor embedded in an automated AVTB air defense system and the CSRDF). External validity refers to the performance equivalence between like entities in the integrated system and some defined "real-world" standard against which the simulation is compared. When appropriate experimental designs are employed, validity tests not only identify differences among systems, but provide a basis for calibrating (modifying) simulations to better achieve performance equivalence.

We categorize procedures for assessing consistency and validity into those that employ experimental techniques that satisfy the scientific criterion of testability and "pretest" procedures that do not satisfy this criterion. Procedures in the latter category include the careful scrutiny of a simulation's modeling process for logical and physical errors (fidelity assessments). The testability criterion requires (1) manipulating multiple factors in factorial experimental designs to permit tests of causal hypotheses about what factors independently and interactively affect outcomes, and (2) using the same experimental designs for the simulation being assessed and the comparison system (another simulation or a "real-world" standard). Pretest procedures—testimonials, face validity, fidelity assessments, and some statistical relationships—can provide important information needed for formulating causal hypotheses. However, these commonly used procedures cannot provide information about underlying *causes* that affect observed outcomes, which is crucial to understanding the internal or external validity associated with a simulation and making correct calibration decisions. When structure is imposed on outcome data by the experimental design, graphic data analyses allow causal trends to be viewed. When the same designs are used to collect data in both simulation and comparison systems, it is possible to pinpoint the location, magnitude, and direction of outcome discrepancies in terms of their underlying causal factors, thus guiding calibration efforts.

When human decisions or judgments are the validity focus, hypotheses address both *what* factors affect the judgments or decisions and *how* they affect them. The *what* hypotheses can be adequately *tested* using factorial designs that reveal if factors affect users' judgments or decisions the same in simulation and comparison systems. The *how* hypotheses require more extensive experimental design features that allow *tests* of algebraic theories to explain judgments or decisions, and *measure* the causal effects. Algebraic models that do not pass their tests are *rejected* along with their measures of the causal factors. Judgment and decisionmaking models that receive empirical support for their validity can (1) measure the magnitude of observed differences in judgments and decisions in simulation and comparison systems, and (2) be directly embedded in a simulation to provide decisionmaking and other subjective inputs with known validity.

The experimental design concepts and procedures described in the text can provide the basis for determining and perhaps achieving high levels of internal validity in the CSRDF/AVTB integrated system. A continuing effort should be pursued to include all critical elements of the system. Since external validity assessments depend on systems external to the CSRDF/AVTB integrated system, they would have to be conducted to the extent possible within cost, availability of credible standards, and other resource constraints.

Proof-of-Principle Recommendation

Based on our assessment of benefit and feasibility, we recommend that a proof-of-principle demonstration be undertaken to integrate the CSRDF with AVTB to determine if, indeed, such an integration could be accomplished within responsible time and cost constraints and if the final integrated system could provide the quality of analytical support anticipated. The proof-of-principle demonstration would have the following objectives.

- Establish the feasibility of integrating a high-fidelity R&D simulator with a SIMNET-based DIS environment capable of supporting weapon system analysis within reasonable time and cost constraints.
- Develop integration tools and procedures that will transfer to future similar integrations.
- Upgrade internal and external validity properties of selected AVTB/SIMNET representations to meet weapon system analysis requirements.
- Demonstrate the breadth and flexibility of a DIS-based analysis.
- Identify additional modeling and simulation validity shortfalls for future upgrade.

Establishing the feasibility of integrating these disparately developed systems in such a way that the resulting system significantly enhances existing analytical tools would open the door to a new era of weapon system analysis. It should be noted that, while the integration itself poses the most difficult technological problems, doing it in such a way as to produce a high-quality analysis capability is a fundamental purpose. Merely integrating the systems in a connective sense is not sufficient in either execution or result.[3]

Throughout the proof-of-principle integration, activities would be conducted in such a way as to develop tools and procedures that would form a basis for integrating disparate simulators with SIMNET in general. Examples would include database interchange, consistency and validity evaluations, protocols, and gateway architecture. If the work is successful in this regard, future integrations should require significantly less preparatory intellectualizing and

[3]Preliminary connectivity was established between the CSRDF and SIMNET for demonstration purposes in early 1993. The CSRDF program architecture was modified for asynchronous processing and data interchange with network elements; communication interface units were developed; and a replica of the CSRDF fixed terrain base was installed on SIMNET. Although this was a commendable first step, it achieved its demonstration purposes without addressing important critical technical issues discussed in this report (e.g., fidelity, consistency, correlation, and validation) associated with weapon system analysis, nor did it provide a tool kit for future integrations.

development of supporting integration software and hardware systems. A "how to" manual would be a useful product from this effort.

In sum, we feel that we have identified in this research a feasible approach to developing a simulation environment that can assess the performance capability of advanced weapon systems in an operational AirLand battlefield environment. We believe that, following necessary internal and external validity work, this simulation environment will provide the Department of Defense (DoD) with a greatly improved tool for analyzing the operational effectiveness of weapon systems. It should also prove effective in later phases of the acquisition process—prototyping and live testing—when choices can be made based on simulated hardware before prototypes are built, and operational tests can be designed without environmental and safety constraints making them more relevant to combat operations. We believe this is a unique opportunity for advancing the state-of-the-art of weapon system analysis and, if successfully accomplished, can provide the weapon system acquisition and development process with more realistic and credible analyses than is possible with present analytic tools, as well as information that spans a broader range of effectiveness questions, thus potentially greatly improving the capability of the DoD to assess future weapon system capabilities.

Acknowledgments

The research in this report resulted from the combined effort of a number of highly qualified people, each with his own special area of expertise. Terry Gossett, Edward Huff, and Lieutenant Colonel James Voorhees of the U.S. Army Advanced Systems Research Office (ASRO) and NASA (Ames Research Laboratory), provided the in-depth knowledge of the Crew Station Research and Development Facility (CSRDF) rotorcraft simulator, all being principals in its development and operation. They wrote the detailed description of CSRDF found in Appendix B. Duncan Miller and Brian Soderberg of BBN Advanced Simulation Division provided the in-depth knowledge of SIMNET and, likewise, were principal developers of that system. Duncan Miller wrote the detailed description of SIMNET presented in Appendix C and contributed primary ideas on applying a networked system to the analysis task. Steven Koons of SYSCON Corporation, under contract through ASRO, developed the detailed integration approach adopted by the study.

Giles Smith and Jeffrey Drezner of RAND served as the experts on the weapon system development and acquisition process.

The authors are most grateful and indebted to all of these individuals for their inputs, ideas, and support. However, we alone are responsible for the final written product and bear the burden of any errors of commission or omission.

1. Introduction

The need for improved methods to aid in decisions concerning the development and acquisition of modern weapon systems becomes critical as weapon systems increase in technological complexity, cost, time to develop and acquire, and uncertainty of performance. These trends could have particularly deleterious effects on the operational capability of the nation's military forces because of a more limited ability to recover from program or design deficiencies in a budget constrained era. Thus, it is necessary to improve our ability to properly assess weapon system performance, beginning with the concepts and engineering designs and continuing throughout a weapon system's development, to identify the exceptionally good concepts and uncover serious design flaws and operational shortcomings so that system or program modifications can be made or poor decisions reversed without incurring high cost penalties. To this end, attention has turned to expanding the application of modeling and simulation in weapon system analyses. As computer power and efficiency have advanced at an unprecedented rate, so has simulation and simulator technology and, hence, the potential for major improvements in weapon system analysis. The brief review of the current weapon system development and acquisition procedure presented next brings out where and how expanding modeling and simulation could have an important effect.

The Weapon System Development and Acquisition Process

The weapon system development and acquisition process for major defense acquisition programs[1] is divided into five phases. Each phase is preceded by a decision point, referred to as a milestone, where a Defense Acquisition Board and the Secretary of Defense determine if and how the program should proceed. Prior to each decision point, there is a thorough review of program status. A program moves through each phase at its own pace. The phases are tailored to the specific program to minimize acquisition time and life-cycle costs and are consistent with the urgency of need, degree of technical risk, and progress as demonstrated by test results. Of primary interest to this study are Phases 0, I,

[1]See Appendix A for a more complete description of the Department of Defense weapon system acquisition process.

and II described in Table 1.1, which lead to a Milestone III decision on authorizing full-scale production at the end of Phase II.

The Analytical Situation

When mission capability requirements are central to initiating the development of a new system, questions of operational effectiveness on the battlefield are of prime concern. Specific questions are

- What effectiveness is needed?
- What effectiveness does the upgraded or replaced system have?
- Why and by how much does it fall short of mission needs?
- What features does the new system need that would lead to greater effectiveness?
- What is the magnitude of the increased effectiveness that would result from those features and is it sufficient to warrant development?

Each milestone decision addresses these questions. It is important that the weapon system development and acquisition process have solid analytical

Table 1.1

Weapon System Acquisition Process

Milestone 0	Decision to approve a mission need and entry into the concept exploration/definition phase.
Phase 0 activities	Explore system design concepts; develop an initial operational concept; and conduct tradeoff studies and risk analyses/assessments.
Milestone I	Decision to approve proceeding into the concept demonstration/validation (dem/val) phase.
Phase I (dem/val) activities	Prototyping; competitive demonstrations in operational environments among prototypes produced by different contractors; finalizing the operational concept; developing the test and evaluation plan; and translating functional requirements into technical design specifications.
Milestone II	Decision to approve proceeding into weapon system development.
Phase II activities	Developing and producing limited quantities of the weapon system; conducting a Development Test and Evaluation; conducting an initial Operational Test and Evaluation; and planning for production.

support to answer the above questions so that, to the degree possible, correct design decisions (i.e., those that result in high operational effectiveness with minimum cost) can be made well before the system is actually built. Although design and production deficiencies may never be completely avoided, their occurrence can be reduced through the persistent use of credible analyses. The use of such analyses should start prior to Milestone 0 to support the decision to approve a mission need, and continue to be used to determine the best possible concept for satisfying operational needs, identifying design shortfalls and preferred design attributes early on, and, in the later stages, confirming that engineering specifications are met and measuring resultant operational capability.

The analytic tools used in the weapon system development and acquisition process today consist of professional judgments, combat simulations, research and development (R&D) human-operated simulators, and prototypes in live tests and evaluations. Prior to Milestone 0, a mission need generally emerges from professional judgments about current capability and projected future threat, developments perceived for that future time period, and other factors relating to sustaining a fighting edge over potential adversaries in potential conflict environments. From this, a case is made to obtain approval to proceed into Phase 0 to clarify deficiencies and explore system concepts and performance requirements that overcome them. During Phase 0, professional judgment continues to play a major role in gauging operational effectiveness potential. Judgments are often based on engineering performance calculations associated with systems (e.g., vehicle dynamics), system components (e.g., sensor system detection), and subsystems (e.g., automated target recognition). Occasionally, results from combat simulations are used in Phase 0 as input to judgments and decisions.

Results from constructive computer simulations[2] are used extensively in Phase I, since by this time a system's attributes have been better specified. During this phase, competing industry teams develop R&D simulators to investigate human/system interfaces and system design issues, and converge technical capabilities with the system specifications. Also, prototypes ranging from brassboard components to full-scale weapon systems that could be used for limited operational effectiveness analyses are developed in Phase I. The results of these effectiveness analyses become an important part of the Milestone II full-scale development decision. In Phase II, the weapon system is actually produced

---

[2]Constructive simulations are those in which all weapon system interactions, force movements, and command and control decisions are automated, and the simulation runs without regard to the actual time required for events to occur (e.g., JANUS, CARMONET, TACBRAWLER).

and analyses focus on live tests and evaluations to assess conformation to engineering specifications and operational needs. These tests can be supported by a high-performance R&D simulator, which can now be calibrated with actual data.

The primary element missing from current analytic tools is the *combined abilities* to

- simulate a weapon system's performance specifications with sufficient validity to measure performance differences in alternative approaches or technologies,
- simulate the weapon system's performance in the military unit configuration and operational environment for which it is being developed, and
- have human operators operate the weapon system in its intended combat environment, making real-time decisions.

Professional judgments, of course, lack all of these features. Combat computer simulations lack the third feature. Human-operated R&D simulators lack the second feature because of the limited size of their combat environments. Prototype live tests and evaluations lack elements of both the first and second features, and prototype components often critical to the analysis (e.g., sensor systems) are missing or consist of simplistic representations. Further, tactics used in operating the prototype weapon systems in operational tests have to be curbed for safety reasons, and combat environments are necessarily limited.

Focus of This Report

An important advancement in operational analyses is evolving with the development of distributed interactive simulator networking. The Advanced Research Projects Agency (ARPA) and the Army jointly developed a simulator network featuring distributed processing techniques that allow large numbers of manned simulators and unmanned simulations of weapon systems to interact on a single simulated battlefield. The system, called SIMNET, provides the basic mechanisms for constructing human interactive environments in which to conduct weapon system analysis.[3] Meanwhile, defense industries and the military service laboratories now routinely develop human-operated simulators

[3] Since SIMNET's inception, other distributed interactive simulation (DIS) networks have been and are being developed that extend/improve on SIMNET. Examples are the Army's Battlefield Distributed Simulation-Developmental (BDS-D) and ARPA's ODIN and Warbreaker. Here we will use the terms SIMNET-based or simply SIMNET to refer generally to such simulator networks.

(called R&D simulators) to test design concepts, investigate human/system interfaces, and conduct other research and development activities on specific emerging weapon systems.

We propose that by integrating these two capabilities, it should be possible to analyze emerging weapon systems in a large human interactive combat environment, combining all three of the analytic features listed above. Such an environment could be of great value in aiding the weapon system acquisition and development decision process. We further propose and describe how to approach an initial integration that would both serve as a proof-of-principle experiment and provide an intellectual, hardware, and software tool kit for future integrations.

We describe in the following sections how an integrated SIMNET/R&D simulator system suitable for weapon system analysis could be achieved and discuss the potential contribution it can make in improving decisionmaking and system design. Section 2 details SIMNET's features, describes characteristics that contribute importantly to weapon system analysis, and discusses integrating an R&D simulator into the SIMNET-based Aviation Test Bed (AVTB) as a separate node in that system. Section 3 raises the technical issues that must be dealt with in such an integration. Section 4 discusses the critical area of modeling and simulation (M&S) validation and addresses issues specific to a SIMNET/R&D simulator integration. In Section 5, we describe a proof-of-principle experiment designed to verify the concept and pave the way for broad exploitation of the approach.

2. Modeling and Simulation Issues and Opportunities

As the primary tool for the comprehensive analysis of the operational effectiveness of developing weapon systems, modeling and simulation (M&S) provides both challenges and opportunities to address the important questions posed in the weapon system development and acquisition process. Simulator networking, in particular, may provide many of the analytic features needed to improve the credibility of analyses used in this process. First, we turn our attention to analytic features needed for weapon system analysis and then describe features inherent in a simulator networking system.

Analysis Features Important to Weapon System Analysis

Operational Analysis Scope

Operational analysis of a weapon system must be conducted over a range of combat perspectives, as shown in Figure 2.1. At the most detailed level, the characteristics of the weapon system's components and those of the threat and combat environment must be identified and modeled (see the activities listed under "tactical-technical" in Figure 2.1). The results form the basis for and provide inputs into high-resolution M&S, where system-on-system interaction is simulated to determine, among other things, the weapon system's lethality, survivability, and overall operational performance on the battlefield. Typically, for AirLand battle systems, this simulation represents combat vignettes from the division perspective, detailing battalion/brigade battles and employing the weapon system in a particular role (anti-armor, reconnaissance, or interdiction). The results of the high-resolution simulations can also provide the means for "tuning" the lower-resolution models that address other important facets of the weapon system. A corps-level simulation aims to assess the weapon system's flexibility in an environment where it must perform multiple types of missions across a larger geographical area for a longer time period than the high-resolution simulation, and considers corps assets as well as cross-divisional employment. At this level, weapon system availability—a function of reliability, maintainability, and survivability—is also measured. Informed by both the high-

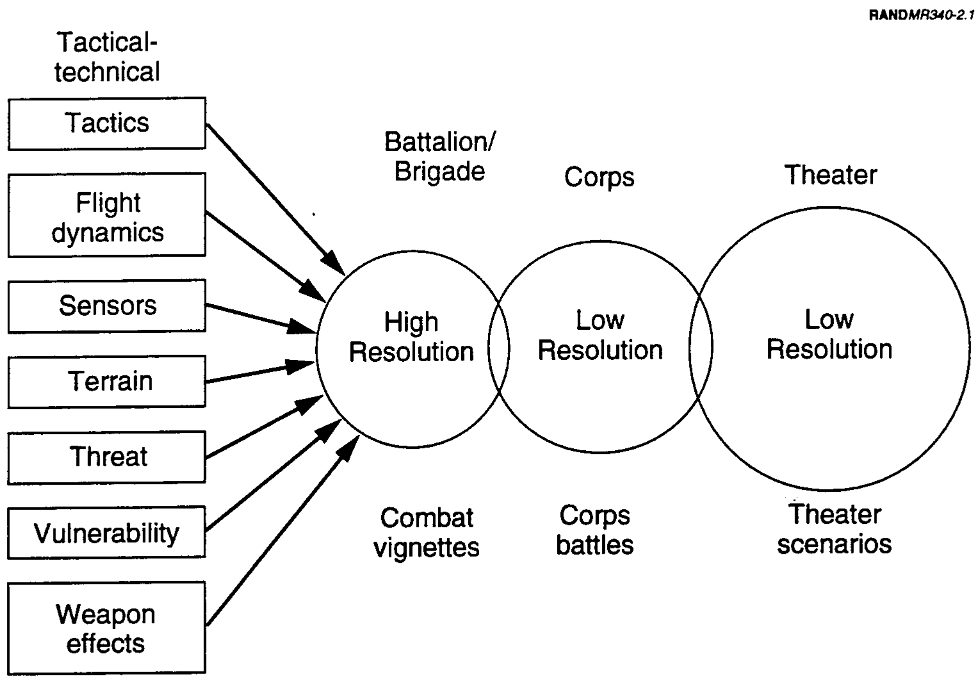

Figure 2.1—AirLand Battle Weapon System Analysis

resolution and the low-resolution corps model results, the theater-level analysis considers the weapon system's contribution to an even broader set of missions, including possibly interdiction, cross-corps, joint, and combined (i.e., with other nations' forces) operations, and a geographically large conflict arena.

Four important elements of combat analyses are shown in Figure 2.2. Operational environments are the terrain, foliage, and cultural features of the operating area; the friendly and enemy forces, their laydowns, battle plans, doctrines, and tactics; weather and other obscurations; and the scenarios that form the backdrop of an analysis. Weapon system performance refers to the physics and engineering attributes that provide and control the physical activity and phenomenology of the systems. Weapon system operation refers to how that performance is actually employed by a human operator. Combat interactions are the result of the confluence of the other three—the seeing or not seeing, engaging or not engaging, winning or losing.

The elements of weapon system analysis (Figure 2.2) and the high- to low-resolution modeling of increasingly more global combat (Figure 2.1) are now commonly incorporated in simulations wherein all systems and actions are fully automated by a computer program.

RANDMR340-2.2

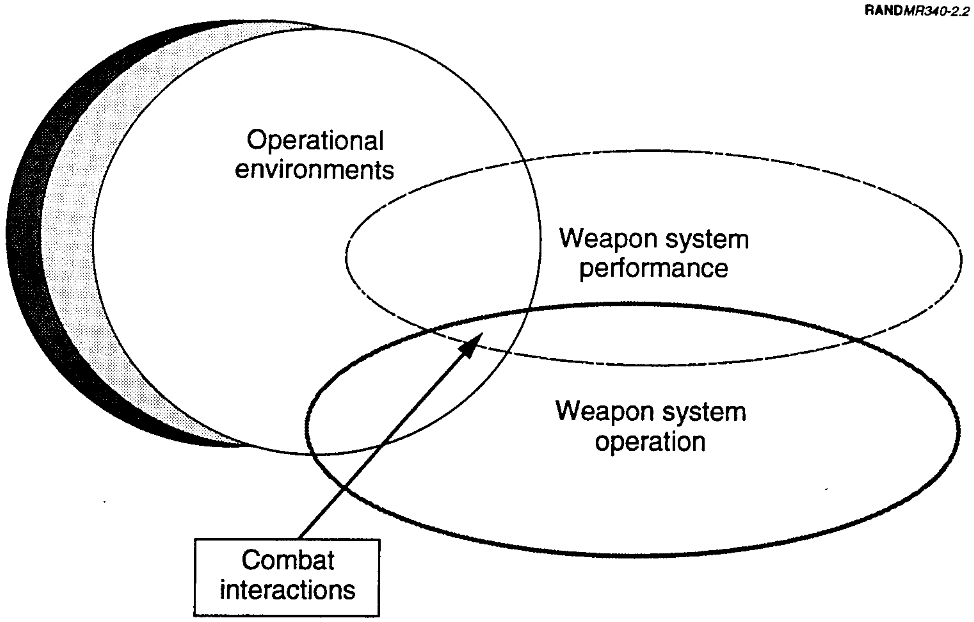

Figure 2.2—Elements of Weapon System Analysis

The Critical Issue of Credibility

The credibility of a simulation can be assessed from two conceptual viewpoints: fidelity and validity. Both concepts refer to the "realness" or accuracy of a simulation's representations along dimensions related to its intended purpose; that is, how closely these important representations approximate some "real world" counterpart. We distinguish between the two concepts based on how assessments of "reality" or accuracy are made. In this report, assessments of *fidelity* result from close scrutiny of a simulation's modeling process for logical and physical "misrepresentations," such as operationally irrational actions (e.g., helicopters hovering over enemy tanks), violations of natural law (e.g., tanks driving on lakes, aircraft flying through trees), incorrect representations of human processes or systems performance (e.g., identical detection probabilities for all sensor-target pairings, instant decisionmaking or action taking), and important errors of omission, to name just a few. Simulations that contain these undesirable "errors" along dimensions important to questions the simulation will be used to address are said to have "low fidelity." Simulations that appear "real," that is, appear to faithfully represent reality given their intended purpose when closely scrutinized, are said to have "high fidelity." Some simulations may have mixed fidelity—high on some important dimensions, low to medium on others. For *validation,* on the other hand, we impose the requirement of hypothesis formulation and testing rather than simply "appearance" for assessing the "realness" of the simulation along perceptual or operational dimensions of concern. Procedures for assessing a simulation's validity and making decisions about upgrading a simulation's credibility are the topic of Section 3.

Simulator Networking

Simulator networking is a concept wherein a large number of manned and unmanned simulators are interconnected through a communications network so they can interact in real time. The potential for applying simulator networking to greatly enhance the analysis methodology for the weapon system development and acquisition process derives from the critical influence of the human operator on system performance and the impact of interactions among human operators on broader operational effectiveness measures. With simulator networking, simulators of the developing weapon system can be assessed on a simulated battlefield as they directly interact with a full spectrum of manned and unmanned threat system simulators. Tactics are continually assessed and improved by human operators, combat interactions reflect the higher motivations and frailties of humans under stress, and the physics and engineering attributes of the equipment are enmeshed with the humans' ability to apply them.

The SIMNET System

The technology underlying simulator networking, developed within the last decade, was initially implemented as the SIMNET system. SIMNET was sponsored by the Advanced Research Projects Agency and developed in conjunction with the U.S. Army. The goal of this program was to develop the technology to network large numbers of interactive combat vehicle simulators and their supporting elements. In effect, SIMNET provides a simulated world in which fully manned platoon-, company-, and battalion-level units can fight force-on-force engagements against an opposing unit of similar composition. The system provides a joint, combined arms environment with a range of command and control and combat service support elements essential to actual military operations. All of the elements that can affect the outcome of a battle are represented, and victory is likely to go to the unit that is better able to plan, orchestrate, and execute its tactical operations.[1]

SIMNET has an object-oriented simulation architecture wherein there is no central "host" computer for event scheduling and master control of the simulation. Rather, each simulation node (e.g., tank or helicopter automated force control station) is autonomous and responsible for (1) maintaining its own state, (2) sending messages on an open communications net describing changes in its state, and (3) interpreting and responding to messages sent by other nodes. Each simulator has its own copy of the nonchanging world—the fixed terrain and cultural feature data base. As the SIMNET network expands, each new simulation node brings its own computational resources and hence has no impact on the other nodes' processing requirements.

Advantages and Limitations of the Current SIMNET System

Simulation networking, in general, has several unique features to contribute to weapon system analysis. There is no *technical* limit on number or type of interacting manned weapon system simulators and semi-automated forces with the potential for quasi-intelligent behavior interacting directly with manned simulators; a semi-automated weapon system is visually indistinguishable from its manned counterpart on the battlefield. This allows for the creation of large-scale battlefield environments on which to perform analyses. SIMNET includes command, control, and communications elements that interact on-line with simulated forces. And it contains an extensive set of analytic tools, including a "stealth" vehicle that can observe the battle from any point on or above the battle (e.g., looking out the windscreen of the cockpit of an aircraft) and a complete playback capability for all or part of the simulation.

The distributed simulation environment pioneered by SIMNET provides the first practical way of generating realistic loadings on all of the battlefield systems that must be taken into consideration. However, since the primary goals in developing SIMNET were to demonstrate the simulator networking concept and implement it as a training system, emphasis was not on producing "realism" of the SIMNET components. For example, image generation, vehicle dynamics, sensor system performance and operations, and other elements were deliberately

[1]See Appendix B for a detailed description of SIMNET. SIMNET and SIMNET derivatives remain the only functional implementations of simulation networking and provide the foundation for the extensive concept and technological development activities in distributed interactive simulation (DIS) currently under way within the defense modeling and simulation community. For simplicity, we use the term SIMNET without deprecation to other named versions such as the Army's Battlefield Distributed Simulation-Developmental (BDS-D) and ARPA's ODIN and Warbreaker systems.

modeled to maintain fast processing rates using moderate capability and relatively inexpensive computers. Also, SIMNET relies heavily on decision tables to determine what view the operator will have of the battlefield. These features were deemed appropriate for the Army's initial training objectives, and the Army considers the resultant system to be effective in achieving its training goals and, in many ways, superior to alternative training methods. However, these moderate to low levels of fidelity may not be adequate to support weapon system development and acquisition decisions.

A High-Fidelity Weapon System Simulator for SIMNET Integration

There are two approaches to achieving a high-fidelity representation of a weapon system in a SIMNET environment. One is to build a special high-fidelity SIMNET model of the system. The other is to integrate into SIMNET the weapons system's high-fidelity R&D simulator(s) normally produced in the development process either by the government or industry—*essentially adding it as another node* on a SIMNET network. The first approach could be very expensive and time-consuming if many weapon system candidates were to be analyzed, and would duplicate the R&D simulators already developed. The second would require major modification to the architecture of the typical existing R&D simulator but might very well provide the best representation, especially as the weapon system design becomes finely tuned in later development phases. Once the principle of analyzing weapon systems in an integrated manner with a simulator network is established, the R&D simulators could be designed at the outset for that integration.

We address the second approach—integrating an existing high-fidelity R&D simulator of the weapon system being developed within the context of a simulator network. In particular, we consider here the exploratory integration of a high-fidelity weapon system simulator with SIMNET and appropriate upgrades to SIMNET components as an important next step in the evolution of analytical techniques to support the weapon system development and acquisition process. The SIMNET implementation we have chosen is the Aviation Test Bed (AVTB) located at the Army Aviation Center at Fort Rucker, Alabama, and the R&D system to be integrated is the Army/NASA Crew Station Research and Development Facility (CSRDF) advanced helicopter simulator located at the Ames Research Laboratory, California.

It is important to note that *integration* for this purpose requires a full scope of interactions among the integrated systems that is logical and appropriate for their physical properties and operational capabilities and the combat environment. This requires careful attention to several simulation features including fixed database interchange; timely, complete, and accurate network communications; precision of terrain features and system locations and aspects; consistency of sensor perceptions; calibration of like-system performance; and validation of the distributed system's characteristics that are relevant to analysis objectives (the last three categories are discussed in more detail in Section 4).

Crew Station Research and Development Facility

The CSRDF developed and operated by the Army and NASA at the Ames Research Center seems ideal for the exploratory development work because of its government ownership, representativeness of complete weapon system simulators, and dedication to combining research with practical application. The CSRDF resulted from a high-priority joint NASA-Army development project. It serves as a full combat mission rotorcraft simulator with emphasis on research into mission equipment and crew vehicle interfaces for rotorcraft weapon systems. The basic combat environment designed for the CSRDF is configured to support a wide range of environments. The crew stations of the CSRDF have glass cockpits that appear to faithfully represent all crew station interfaces for daytime or night adverse weather missions. All displays in the cockpit are reconfigurable, and the pilot has a fiber optic helmet-mounted display (FOHMD) that provides a 67° vertical by 127° horizontal field-of-view and an unlimited field-of-regard. All subsystems appear to have the appropriate levels of interaction with other subsystems and combatants (e.g., fields of view, intervisibility, and countermeasures at all wavelengths are modeled). The CSRDF uses General Electric's CompuScene IV image generator.

The CSRDF provides limited capabilities for combat simulation and evaluation. A maximum of 100 ground elements and 11 rotorcraft elements can be simulated. The air elements include a Scout/Attack team, utility rotorcraft, or threat rotorcraft. All ground elements are automated and scripted, whereas air elements are controlled through the use of three interactive graphics workstations that provide the capability to control up to four aircraft each. The CSRDF is a traditional simulator consisting of a central computer complex that drives the research cockpits and auxiliary workstations that provide simulated interaction with the research cockpits.[2]

[2]See Appendix C for a more detailed description of the CSRDF.

Mixed-Fidelity Simulations for Weapon System Analysis

Large-scale, mixed-fidelity, man-in-the-loop simulation networks appear to provide a promising way of performing credible operational analysis and setting the stage for upgrading the weapon system development and acquisition process.

High-fidelity R&D simulators could play a crucial role. In the early decisionmaking phases, the system being evaluated has not yet been built and has no existing real-world analog, so it would be necessary to use detailed simulations that closely represent important underlying physical phenomena (detailed in system specifications) to predict system capabilities and operational characteristics. These capabilities should account for crew performance and hardware characteristics and limitations. The CSRDF is a good example of a simulator designed to conduct these sorts of studies. Various modules can be incorporated to represent airframe modifications and sensors. Crew performance can be measured and studied in considerable detail. Navigation and communication systems are modeled and employed during these studies so that cognitive and psychomotor loads on the crew members are realistic. Finely detailed representations of underlying physical phenomena are used to generate models of sensor performance or of airframe stability and handling characteristics. Once these models have been generated and the dynamic behavior of the system has been explored, it would be possible to replace these physics-based models with simpler ones that exhibit similar behavior. In most cases, the simulation host computer is not capable of running all of the highest-detail modules simultaneously. Once the crew/vehicle/sensor model's parameters have been established, it becomes possible to expand the simulation context to include interactions with other man-in-the-loop simulations. Here is where lower-fidelity simulators can play an important role.

Lower-fidelity simulations can provide a more realistic battlefield for studying small-unit dynamics and tactical behavior. This allows focus on interactions among various simulated entities rather than on a single simulator. The parameters of the lower-fidelity simulators should be adjusted so that their externally observable behavior corresponds closely to that of the higher-fidelity manned simulators. Externally observable behavior would include operational performance envelopes (maximum and minimum velocities, longitudinal and lateral accelerations, etc.), sensor ranges, weapon effectiveness, electronic countermeasure effectiveness, and so forth. Key pilot/vehicle and pilot/sensor performance should be calibrated as well, to ensure that target detections and hit probabilities in the lower-fidelity simulators are similar to those obtained in the higher-fidelity simulators. Tactical team behavior could then be studied in a variety of operational scenarios, and models formed for inclusion in semi-automated simulations.

Semi-automated simulations are critical to the simulation of large-scale battlefield operations, since it will rarely be possible to assemble the large numbers of manned vehicle simulators and crews needed to assess battlefield operations at the battalion/regimental level and beyond. This is especially the case when repeated runs are envisioned to explore variations of a system configuration and an operational scenario. To avoid the complexities raised by learning effects (that would occur in the manned simulators), new populations of crew members would continually need to be recruited and trained, an often prohibitively expensive and time-consuming activity. Semi-automated forces, on the other hand, can be programmed to execute well-defined behaviors repeatedly without tiring or learning. This feature is both a strength and a weakness, of course, and is the fundamental reason why a mixed high-fidelity/lower-fidelity/semi-automated simulation strategy might be beneficial.

Analysis Potential from an Integrated High-Fidelity R&D Simulator and SIMNET Environment

There appear to be a number of potential significant improvements in our ability to support weapon system development and acquisition decisionmaking by integrating an R&D simulator into a SIMNET environment. For conducting weapon system operational effectiveness analysis, a system consisting of a high-fidelity R&D simulator integrated into a SIMNET environment could provide a greatly improved capability over current combat simulations and an improved capability over SIMNET's lower-fidelity simulators. The weapon system analysis schematic shown in Figure 2.1 would change to that shown in Figure 2.3. By replacing a constructive simulation at the high-resolution level with either fully manned or mixed manned/semi-automated forces on the battlefield and adding real-time command and control, the enhancements would include those shown in Table 2.1. These enhancements feature real-time interactions among combatants and evolution of tactics and techniques throughout the course of the analysis.

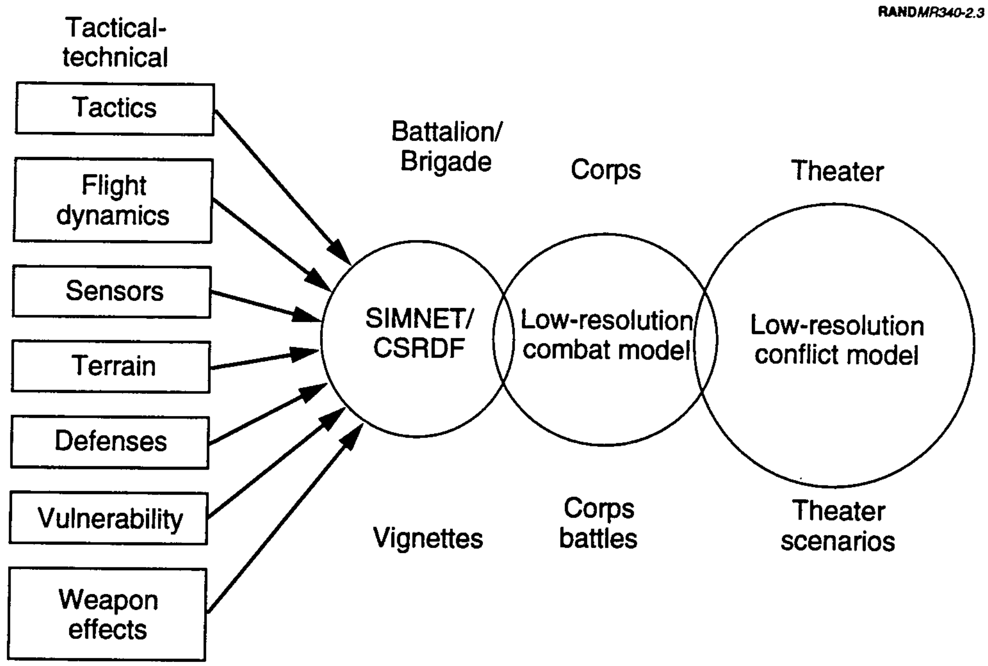

Figure 2.3—Weapon System Analysis Potential with SIMNET/CSRDF

Table 2.1

Potential Analysis Enhancements

Environment	Current	SIMNET/R&D
Tactics	Specified, adjusted	Evolved during analysis
Vehicle	Coarse dynamics	Crew interaction
Sensor models	Varying fidelity, notional operability	Varying fidelity, crew operability
Terrain	Contours only	Contours, foliage, and culture
Threat	Variable fidelity, notional operability	Variable fidelity, crew operability
Vulnerability	Engineering estimates	Engineering estimates
Weapon effects	Flyout/statistical	Flyout/statistical
Combat interactions	Stylized	Real-time, human-in-the-loop

The contributions shown in Table 2.2 to operational test and evaluation (OT&E) conducted in Phase II full-scale development would follow from the ability to conduct pretest assessments to identify the most important areas to investigate, determine deficiencies in OT&E designs and implementations, and train personnel to conduct the tests and evaluations. In concept, pretests could be run using a replica of the test environment, including instrumentation and manned weapon systems. The simulator network would also permit investigation of performance outside the safety envelopes required for operating actual systems. More complete combat environments could be simulated than could be practically and perhaps feasibly established on the test range, and advanced future capabilities that do not exist yet could be included in the assessment.

In the next section, we address the technical problems that must be overcome to achieve an integrated simulator networking system capable of supporting weapon system analyses.

Table 2.2

Potential Contribution to Operational Test and Evaluation

- OT&E plan assessment
- Early operational assessment
  - Identification of critical issues
- Pretest investigation
  - Assess and hone test design
  - Familiarize participants
  - Avoid execution flaws
- Compensate for real-world limitations
  - Safety factors (tactics, weather, night)
  - Multiple weapon systems employment
  - Density of threats and targets

3. Major Technical Issues in the Integration of AVTB and CSRDF

SIMNET-based systems and R&D simulators represent different design philosophies and architectures, giving rise to a number of technical issues that must be resolved to make the AVTB/CSRDF integration successful. They include

- coordination of synchronous and asynchronous processing
- simulation network communication and protocols
- database sharing and correlation
- sensor modeling and visual image representation
- correlation between simulator modes.

In this section of the report, we provide an overview of the major issues that must be considered in integrating a flight simulator into AVTB. Appendix D gives a more complete analysis of the details involved in integrating the CSRDF with AVTB.

Coordination of Synchronous and Asynchronous Processing

AVTB and CSRDF have different architectures. AVTB processing is distributed and asynchronous. No master computer exists in the system to calculate vehicle kinematics, sensor performance, and scenario interactions. Each simulator has its own processor that is responsible for calculating local vehicle interactions based upon data received from other simulators. Data are transferred in an asynchronous manner between simulators on the local area network (LAN) when some specific change occurs in the state of the simulated vehicle (e.g., heading or speed change, weapon firing, etc.). The CSRDF, on the other hand, is a synchronous system with a single host computer. Computations and data transmissions between elements are keyed to a basic clock or data transmission cycle. Thus, for example, data transmissions between the CSRDF host computer and its team stations occur on a periodic basis with data consisting of database elements required to display scenario interactions at the team stations. When the CSRDF resides in the AVTB environment, it will have to react to incoming AVTB messages about the status of other vehicles and their actions and appearances, and interrupt its own synchronous cycle, determine how and when to display the transmitted events, and react with its own modeled systems. Resulting time and space inconsistencies between the AVTB and CSRDF must be resolved before this can occur.

Simulation Network Communication and Protocol Issues

Special protocols were developed for data transmission among simulators that are part of SIMNET-based systems, whereas CSRDF uses relatively standard protocols for data transmission between the host computer complex and the team stations. The differing characteristics of AVTB and CSRDF mean that the interface between these two systems must provide a gateway function. A generalized solution for interfacing the two systems could be implemented in one of two ways. First, the hardware and computer programs necessary to interface a high-fidelity simulator to the SIMNET environment can be installed in the host processor of the simulator. Second, a bridge processor can be installed that transforms data from the high-fidelity simulator format to the SIMNET format. The second option is illustrated in Figure 3.1 and appears to be the option of choice to enable a R&D simulator such as the CSRDF to maintain its integrity as an independent research simulator.

In this implementation, a bridge system is installed that gathers data from the CSRDF host processor, determines if AVTB update requirements are met, and transmits data on the AVTB network with the proper protocols. Conversely, data received from AVTB is properly integrated for the CSRDF environment and passed on to the CSRDF Common Data Base (CDB). This is the most generalized solution because no specific requirements (other than data) are imposed upon SIMNET simulators, and the basic processing functionality of the simulator is not affected. Further, additional processing demands are not made on a processor attached to SIMNET.

This general solution should suffice for the connection of an arbitrary simulator to SIMNET. That is, any number of internal stations or configurations (single or multiple crew) can be accommodated. However, it is apparent from the data listed in Appendix D that additions to the SIMNET protocol suite must be implemented. The additions would support enhanced weapons processing, sensors, countermeasures, and other features not implemented in the current protocols. Two specific changes required are:

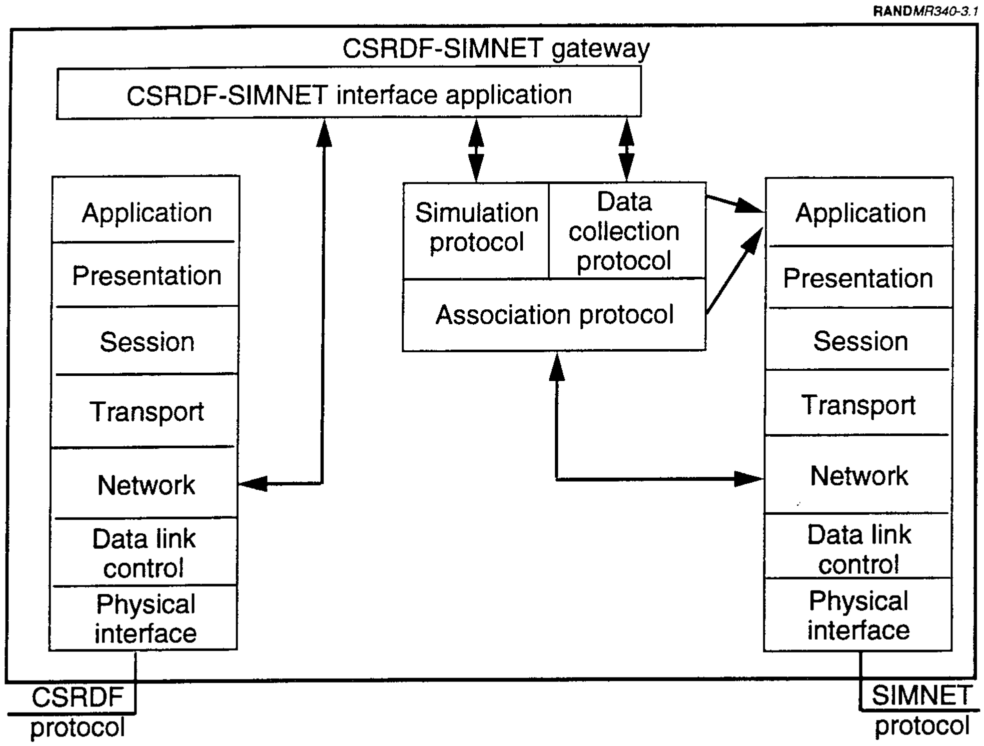

Figure 3.1—CSRDF Bridge Processor Data Interface

- **Voice communication interfaces.** The voice communication capabilities of the CSRDF are substantial. All requirements for tuning, jamming, and voice transmission must be met through the voice communication interfaces for the CSRDF to be properly utilized.
- **Guaranteed message delivery.** Some mechanism must be developed to ensure that messages are delivered in the SIMNET datagram/broadcast environment. Not all messages need be guaranteed delivery, only those without self-healing capabilities such as vehicle kinematics and some sensor phenomena. Messages that fit this category include one-shot laser ranging, weapons fire and detonation, and the like. Any one-time message in which simulation data could be lost is included.

CSRDF models of weapon systems and subsystems (e.g., sensors) typically have many more parameters than those in AVTB, and transmit and receive large amounts of data to function properly. Some of the key areas requiring additional data streams are electronic warfare (EW), EW countermeasures, and command, control, communications, and intelligence (C3I) protocols, scenario control and malfunction messages, and simulation environment control. The CSRDF also has a wide range of communication capabilities that must be supported to ensure that heavy contributors to pilot workload are properly implemented and accounted for.

AVTB itself has many protocols to be interfaced. These include the simulation protocol itself (containing vehicle appearance, weapons fire, collision, logistics resupply, and activation/deactivation elements), the data collection protocol (used for monitoring vehicle status and events of interest to data analysts), and the association protocol (used for establishing and modifying multicast groups for efficient communications). A formidable task in interfacing AVTB and CSRDF is working out the myriad details of exercise initialization, unit assignments and nomenclature, logistics allocations, and so forth. Although formidable, this task is reasonably straightforward. A DoD-sponsored Interoperability Standards Working Group is continuing discussions to work out these details, using the current SIMNET protocols as a basis.

Database Sharing and Correlation

The integration must ensure that databases on the two systems represent the same terrain characteristics, vegetation, roads, bridges, buildings, etc. Disparities may arise from use of different data sources, from storage of different object attributes, or from generation of different representations from the datasets.

Generally, databases will correlate poorly with one another if they are derived from different data sources. Therefore, the integration must utilize a common, machine-independent reference database from which to build the application databases so that both SIMNET and CSRDF will start with the same data. This reference database must include both spatial (geometric) and nonspatial attributes of all features and objects in the database. The geometry of objects establishes their position, size, orientation, and visual appearance, and determines the potential intervisibility between pairs of objects. Nonspatial attributes include texture, brightness, color, and thermal values.

Sensor Modeling and Visual Image Representation

AVTB and CSRDF both generate images by synthesizing and displaying scene polygons. Depending on the display being represented (out-the-window, hard optics, TV, forward-looking infrared [FLIR]), there may be large differences between the two simulation systems in the number and type of polygons generated, although both simulators may faithfully reproduce the salient characteristics of the system. The objective is not to achieve an exact match between reality and simulator-presented image, or between simulators, but to produce a sufficient realism and consistency so the operator behavior and system performance in the simulators can emulate those of the actual system. In general, all players must appear in the same positions and times in different simulators, and cues for detection, tracking, and other activities must be consistent given clear line-of-sight.

Several other problems must be resolved to ensure that correlation and consistency between dissimilar image-generation systems are maintained. For example, special constructs such as stamps and texture must have analogies across systems so that the views presented are appropriate and the performances of system users are equivalent. Where phenomena such as muzzle flashes or dust clouds raised by moving vehicles are critical battlefield cues, it is important that they be represented similarly to all participants. This does not mean that they must appear identical, but they should be equally recognizable to everyone, at similar ranges and decision latencies.

The latter issues primarily involve the visual system and its associated databases. Many of the factors involved can be quantified rather precisely and controlled. Important visual system parameters include image resolution, polygonal processing throughput, level of detail processing, and texture processing. Database parameters include terrain and feature location and sampling resolution, textures, colors, and object attribute values.

Visual Image Correlation

The AVTB and CSRDF simulator visual systems (image generators) will always present somewhat different renditions of the same visual scene. The differences are inherent in the design of the systems and affect various image-generation processes and characteristics such as polygon processing, color, texture, image resolution, and transparency, among others. It also involves accurate reproduction of sensor characteristics such as noise, control settings, and system dynamics. These differences, however, should not preclude such systems being integrated into a networked simulation for the purpose of weapon system evaluation. The question is, which parameters must be correlated and how closely?

The most critical measures are those most directly related to battlefield performance. Target detection and recognition probabilities and hit/kill ratios are particularly important. Among the factors that contribute heavily to these measures are:

- Terrain geometry and the size, position, and orientation of objects on the terrain must be similar to maintain line-of-sight (LOS) consistency and reproducibility of actions.
- The appearance of weapon effects must be recognizable by all participants.
- Fields of view, horizon ranges, and image degradations due to haze and other environmental factors should be sufficiently similar that target detection and engagement behavior is essentially equivalent between the different systems.
- Scene management systems that cull out targets and scene features because of processor limitations should do so in a consistent manner. For example, targets should not appear and disappear indiscriminately between simulations.

The forms of presentation, in terms of the types of polygonal models, sharpness of image resolution, distribution of color and texture, and presence of transparency, haze, and anti-aliasing are also important for visual image correlation. Each of these is discussed in turn below.

Polygonal models. All entities in the database, including the terrain itself, are portrayed as collections of polygons. The shape and complexity of these objects and the level of detail in the database in general are directly related to the number of polygons used in this representation of the world. Computer image generators (CIGs) differ in their capacity to process polygons. The question here is to what extent the polygonal representation of the same terrain and object models can differ before two visual systems are no longer interoperable. As mentioned above, LOS correlation is considered to be of primary importance. An exact correspondence of other properties of structures—shape, detail, etc.—is probably not as critical.

In fact, if polygonal representations need to correspond closely, some simulation parameters can be adjusted to compensate for processing differences. During any given frame, the image generator attempts to process and display all the objects within its pyramidal view volume defined by the angular field-of-view (FOV) and the viewing range. Decreasing the FOV or the viewing range, for example, would reduce the view volume and allow for increased polygon density. We need to determine if this skews the detection or engagement performance in other ways. It is also possible to enhance the detectability of an object by artificially increasing its contrast or detail level, or by reducing the background clutter in the local area. Again, use of this adjustment as a means of compensation must be checked for side effects.

Image generators typically take advantage of the fact that details are less visible at greater viewing distances. They accomplish this by dynamically transitioning object representations to lower level-of-detail (LOD) models (containing fewer polygons) as they move off into the background. The appearance of these less detailed models, as well as the transition distances and methods of transitioning, are other system perceptual differences that may need to be understood.

Image resolution. The sharpness or resolution of the visual image is determined, for the most part, by the number of pixels and the FOV, and is measured in terms of arc-minutes per pixel. Color, texture, and rendering algorithms such as anti-aliasing also affect sharpness. Image resolution is particularly important for target detection and recognition. Visual systems with higher resolutions will provide an apparent advantage over lower-resolution systems to the extent that the task at hand requires the higher resolution. The polygon generation rate of the GE CompuScene IV used in the CSRDF is approximately five times that of SIMNET's GT 110 system, and ultimate resolution is about twice that of the GT 110.

Image resolution interacts with FOV for a given visual system. The combination translates to a relationship of arc-minutes per pixel and overall scene extent. These parameters dictate the sharpness and size of the final image. Final image convolution and anti-aliasing are also related to image sharpness, as the final image is blurred or degraded to produce a desired sensor characteristic. This sharpness has an effect on operational measures such as detection and recognition, and therefore must be well understood. While ground vehicles only need to see a few kilometers, with restricted FOVs, problems arise with long-range sensors on board helicopters, fixed-wing aircraft, and air defense units. Wide-FOV out-the-window (OTW) displays are especially troublesome, because the combination of close-in detail needed for nap-of-the-earth (NOE) flight and masking and large FOV cannot be achieved with SIMNET as it is currently configured.

Viewing ranges of sensor visual systems will need to be matched to allow for coordinated tactics. For example, if one pilot can detect a target at 7 km on his sensor displays, while another one can see out to only 3.5 km, an obvious (and often unacceptable) constraint will be placed on target acquisition. In most of our applications, a 7–8 kilometer horizon appears to be the minimum acceptable limit.

Levels of detail in models and terrain come into play in correlating images of two sensors. Multiple LODs for both models and terrain are commonly used in visual systems to efficiently use the limited polygonal throughput of a system.

These LODs, and the transitions among them, must be carefully evaluated, especially for terrain. As was previously noted, LOS correlation is very important in combined arms tactical scenarios. If a tank crew has successfully screened its vehicle from view in high-level-of-detail terrain but appears inappropriately exposed to another crew whose visual system is using a lower level-of-detail representation of the terrain, target detection and destruction become possible in cases where they would be otherwise not permitted. In the worst case, this situation could lead to unrealistic, simulation-induced tactical behavior that might affect the outcome of the evaluation.

Image generators of different resolutions may be integrated into a networked simulation provided they are assigned to simulation tasks appropriate to their capabilities. For example, dismounted infantry may need to be able to recognize vehicles only at 2–3 km in clear conditions, whereas helicopter pilots may have to be able to recognize vehicles at 6–7 km under the same conditions. Other factors such as contrast, context, and movement may lessen or intensify the effects of image resolution differences, as well. The results of this proof-of-principle experiment will allow for a better understanding of the implications of varying image resolutions.

Spectrum, color, and texture. In addition to geometry, color, texture, and spectrum are important to the perception of a visual scene. These differences among image-generation systems are more subtle, and their effects on performance in the simulator are not as well understood. SIMNET has demonstrated that color and texture changes in targets and their surrounding environment affect target acquisition. Of special interest are texture differences in various spectral bands—visible, near-IR, and far-IR. Color and texture are not critical elements in this phase of our study, yet they may in fact affect the results and should be accounted for.

Modeling sensors for the near-IR, far-IR, and visible spectra can be extremely important to accurate simulation, yet these are modeled quite differently in AVTB and CSRDF. Electro-optical sensor visual systems, whether IR or visual, typically have a complex set of distinct functions they perform to transform the sensor database representation to an image or pixel representation, as shown in Figure 3.2. The lowest-level data representation is typically a polygon with luminance or thermal attributes. Simply speaking, this set of three-dimensional points with interconnections describes an element of the terrain or a moving model along with its brightness or temperature value.

RANDMR340-3.2

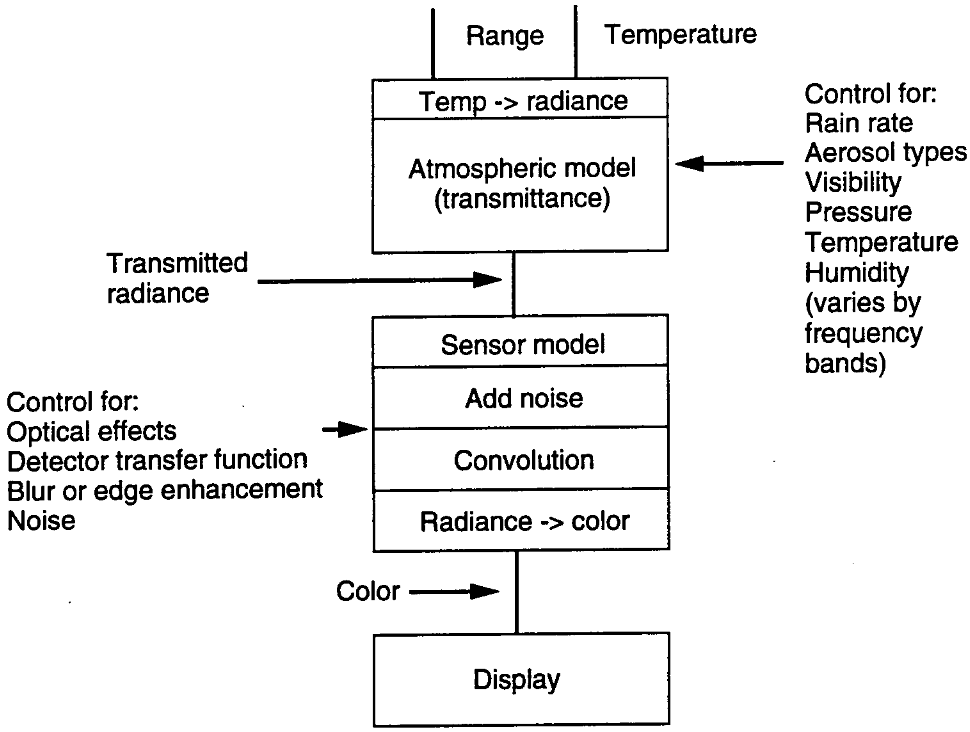

Figure 3.2—CIG Sensor Functional Block Diagram

The first process to take place in a sensor system is the conversion of luminance or temperature to radiance. This process depends on the sensor type. For example, a FLIR system may sense the 8–12 micron band, whereas a television sensor will operate in the 0.4–1.0 micron band. CSRDF performs this transformation explicitly, whereas AVTB simply associates a set of radiances with each pixel of the target icon. This radiance level, along with other parameters such as range from sensor viewpoint and perhaps altitude, is then passed through an atmospheric model. This model incorporates all atmospheric effects such as rain, aerosols, visibility, pressure, temperature, and humidity as they affect various spectral bands. In general, this involves a transmittance table, in which the input radiance value gets converted to a transmitted radiance value. Both SIMNET and CSRDF incorporate atmospheric degradation to some degree.

The third sensor function is an image space convolution to modify the image radiance map to incorporate effects such as blur or edge enhancement, noise, and optical effects. This convolution varies for different sensors such as a low-light television or an 8–12 micron FLIR sensor. Only CSRDF has an explicit algorithm for producing such effects.

The fourth transform is the conversion of radiance level to color and gray level for final image display. Visual sensors involve a flexible mapping of radiance to color and gray levels, whereas thermal sensors just map radiance to gray levels. CSRDF incorporates a dedicated processor to produce such transformations; SIMNET simply displays a preset icon.

These four CIG sensor processes must be represented at approximately similar levels of detail to produce reasonably correlated sensor images. If a lower-fidelity sensor CIG does not incorporate all of these processes, they will have to be added or otherwise accounted for. The CSRDF and AVTB displays can be expected to exhibit large differences in thermal image detection performance.

Transparency, haze, and anti-aliasing. Transparency, haze, and anti-aliasing also affect the appearance of the final rendered image and may influence visual-system interoperability. Trees and bushes, for example, are typically rendered partially transparent so that an object behind them may be seen. Different image generators may use different algorithms to accomplish this effect, resulting in varying levels of obscuration. Most visual systems incorporate a capability for simulating atmospheric haze. Haze blurs the scene and obscures objects in the distance. CSRDF defines a floor and ceiling to the haze function, along with a color shift with distance. Different extinction coefficients may be set for visible, near-IR, and far-IR. The selection of a set of particular mathematical haze functions and their implementation requires careful consideration in striving for interoperable visual systems.

Anti-aliasing algorithms are used in image generators to minimize the apparent stair-stepping effect of pixelated edges. Each pixel edge is blurred with a blend of the colors on each side of it. This technique is particularly useful for lower-resolution systems with larger pixels. The negative effect is that distant objects lose definition, with different effects for different algorithms.

Other phenomena. Special dynamic phenomena such as smoke dispersion, dust clouds, muzzle flashes, missile plumes, flares, and rotor flicker may be produced very differently on the two simulators. Some, such as dust clouds, are currently produced only on SIMNET. Others, such as flares and rotor flicker, are produced only on CSRDF. A common set of dynamic stamps or icons, with similar lifetimes, transparencies, and signatures, should be produced by both systems.

Radar display presentation is handled much differently from electro-optical imaging simulation. The procedure should involve specification of target and background radar cross sections (RCSs), modeling of radar sensitivities in terms of signal-to-noise and signal-to-clutter ratios, determination of contacts, and display presentation. Both threat and friendly radars should be considered, with a range of operating frequencies and conditions. All units should be considered to have some level of Doppler discrimination. CSRDF models radar contacts in this manner, whereas AVTB accesses tables that define probabilities of detection according to range, target aspect, degree of exposure, and presence of movement. It is expected that very different contact phenomena will result from the two simulations, even after adjustments to the table entries. Critical radar operations in AVTB may have to be modeled as they are in CSRDF, including effects of jamming and chaff. Radar warning receivers, now modeled in AVTB, may also have to be modeled explicitly in both systems.

Observations on Sensor and Visual Imaging

During the analysis of SIMNET and CSRDF visual systems, it was noted that neither system provides all phenomenology considered necessary for the use of the simulators in an environment that could affect prototyping and acquisition decisions. These capabilities include

- linking of humidity, diurnal, and atmospheric effects on images,
- thermal distribution and latencies on targets/threats, and
- special environmental factors—fog, smoke, camouflage, etc.

Consistency of imagery begins with the use of a common terrain database interchange specification. Resolution of consistency issues must include a determination that the visual representations across systems do not provide an unfair operational advantage to one player over another.

4. Evaluating the SIMNET/CSRDF Integrated Product

Overview

The credibility of a simulation networking system and an R&D simulator integrated product should be assessed and established before using it for weapon system analysis. A complete assessment process would include verification, validation, and accreditation (VV&A). Verification confirms that the computer code represents what it is intended to represent. Accreditation addresses the simulation's acceptance within the community it is intended to serve. Validation addresses the closeness of a simulation system to a defined standard. The validation process is described and illustrated in this section. Verification should precede the validation process; accreditation may well depend on the results of the validation process.

Two validation aspects of an integrated SIMNET and R&D simulator system would need attention: its *internal validity* (consistency) and its *external validity.* Internal validity refers to performance equivalences among like entities and processes *within* a simulation system (for example, a particular human-operated simulator and its automated counterparts within SIMNET, or like entities between SIMNET and the integrated R&D simulator, as, for example, the CSRDF helicopter simulator and its SIMNET counterparts). Internal validity concepts also include process equivalences (or appropriate *differences*) among different systems using the same subsystem (e.g., detection/engagement decisions resulting from an IR sensor embedded in an automated AVTB air defense system, and an IR sensor as a subsystem of the CSRDF). Internal validity is necessary for simulated operations to occur on any potentially believable level. Whereas internal validity refers to the performance equivalence among like entities *within* a simulation system, external validity refers to the performance equivalence of like entities *between* the simulation system being evaluated and some defined external "real world" standard to which the simulation is being compared. It is important to note that the higher-order purpose of the validation process is to provide a basis for, and guide the way to, achieving the desired validity level, not merely to determine that the desired level has *not* been achieved. This purpose imposes a much stricter demand on the form of the validation process.

Internal Validity Issues Within SIMNET-Based Systems

The validity focus could be on the performance of the hardware itself (in terms of physics or engineering features), how the human operator is represented as using the system in the semi-automated forces, or how the operator actually uses the system's simulator. Within the SIMNET system, internal validity issues concern performance equivalences among like human-operated systems, among like semi-automated forces, and between like human-operated systems and their automated counterparts. They include

- capabilities of weapon systems (e.g., helicopters, fixed-wing aircraft, tanks, air defense units) along their important dimensions,
- capabilities of weapon system components (e.g., ordnance, sensor systems) along their important dimensions, and
- human behavioral representations (e.g., decision rules, flying tactics).

The third bullet refers to subjective judgments and decisionmaking processes that play an important role in operational outcomes. For example, operators interface with simulated sensor systems in detecting and recognizing targets. Detection and recognition capabilities depend upon the physics of the simulated sensors and the operators' perceptions. Also, operators make decisions about weapon system handling tactics based on perceived threat. To achieve internal validity between simulator weapon systems and their automated counterparts, it may be necessary to model important human perceptions and decisions that occur in simulators for incorporation into their automated weapon system counterparts. Procedures for capturing these sorts of human thought processes are described in the modern measurement literature (see, for example, Krantz and Tversky, 1971; Veit, 1978; Anderson, 1981; Birnbaum, 1990); extensions to complex military systems are presented in Veit and Callero (1981) and Veit et al. (1984). The logic underlying these procedures will be presented later in this section.

Some SIMNET inconsistencies (or violations of internal validity) appear obvious at the outset. For example, in the human-operated AVTB helicopter simulators, operators detect, recognize, and classify targets using a simulated TV and IR sensor. However, a look-up table determines target detection, recognition, and classification for the SIMNET automated forces. The look-up table is artificially restricted: only one variable—range of target from the weapon system—changes probabilities. Time for detection to occur is not accounted for in the tables; a weapon system instantly "sees" (or doesn't "see") the target, according to the probabilities entered into the table. Thus, relevant internal validity inquiries would concern determining how the probabilities in the look-up table differ from those produced by an operator in the simulator, and how the look-up table probabilities can be calibrated to more nearly approximate those of the human-operated simulators.

Internal Validity Issues Between the CSRDF and AVTB

Validity issues of consistency between the CSRDF and AVTB concern the following performance equivalences:

- Capabilities of the CSRDF and its AVTB simulator and automated helicopter counterparts,
- Capabilities of CSRDF weapon system components (e.g., sensor systems, ordnance) and their AVTB simulator and automated counterparts,
- Human behavioral representations (e.g., decision rules, weapon system handling tactics),
- The AVTB and CSRDF operational environments in terms of appearance and position of terrain, targets, tree clusters, foliage, roads, rivers, and buildings, and
- The occurrence in time and space of objects and events in the AVTB and CSRDF environments.

An example of an obvious inconsistency between the CSRDF and AVTB human-operated simulators is that the simulated IR sensors produce different quality images. The higher-fidelity CSRDF IR sensor accounts for many more factors, thus producing "fuzzy" images—more like a real IR sensor—whereas the AVTB simulator provides perfect icons to the sensor operator. When such inconsistencies are observed or known to exist, an important question to address is what difference they make in terms of operational outcomes. In this IR example, the specific question is, "How does this difference in image fidelity affect operational outcomes of detection, recognition, and classification, and how can the systems be calibrated so as to reconcile observed differences?" Such obvious inconsistencies need to be investigated and corrected before the system is used for analyses. Internal validity could be improved by calibrating the AVTB IR sensor to the higher-fidelity CSRDF sensor.

External Validity Issues

There will be numerous *external* validity questions associated with the AVTB/CSRDF integrated product. Addressing these questions requires defining a "real-world" standard system (the selection of which is discussed later) against which to compare activities of interest in the AVTB/CSRDF system. Many external validity questions can be constructively investigated simultaneously with internal validity questions to guide the direction of simulation calibrations should take to achieve both internal and external validity. For example, the difference in fidelity between the AVTB and CSRDF IR sensor images could result in large differences in important operational outcomes. If this occurred, external validity data could guide calibrations of one or both of the IR sensor performances in the direction of the validity base (rather than calibrating the lower-fidelity AVTB sensor to that of the CSRDF, as posited above). Particular external validity questions pertaining to the CSRDF/AVTB system include the following:

- Do simulation weapon systems perform like their "real" weapon system counterparts?[1]
- Do operators employ simulated weapon systems and their components in the same way as they employ their "real-world" counterparts?
- Do automated forces affect outcomes in the same way as human-operated "real-world" forces?
- Do algorithms intended to represent human decision or perceptual processes really represent those processes?
- Do humans interface with the simulator environment (databases) as they would interface with its "real-world" counterpart?
- Does the simulation's operational environment—terrain, tree clusters, foliage, weather, and dust—match its "real-world" counterpart in time and space?

Calibrating a Simulation to a Selected Standard

In a practical sense, internal validity is assessed with the idea of calibrating (adjusting) one or both simulations to achieve performance equivalence; external validity is assessed with the idea of calibrating simulations to a defined "real-world" system. In an integrated product, one of the simulations (e.g., the CSRDF) may have greater fidelity than the other system (e.g., AVTB). When this is the case, it might be tempting to designate the higher-fidelity system as the standard against which to internally calibrate the other system. However, examples illustrated later will demonstrate the importance of simultaneously

[1] The term "real" refers to the defined validity standard, the selection of which is discussed later.

using external validity data to guide internal calibration efforts to achieve the best calibrated result, or perhaps even to choose the lower-fidelity system to avoid facing infeasible or impractical modifications associated with extending a lower-fidelity system beyond its basic design limitations.

Internal and external validity data analyzed simultaneously provide information on locus, magnitude, and direction of changes to one or both systems necessary to achieving both validities, when appropriate experimental designs are used (discussed in the next section). After one set of calibrations is implemented, experiments can be repeated to examine results of the calibrations with respect to both internal and external validity. This experiment/calibration feedback loop is continued until the simulation is "acceptable" or economic or practical considerations make it unreasonable to proceed with further experiments. When seemingly unacceptable discrepancies still exist in the final product, a judgment has to be made about the validity of the calibrated simulation. Determining an "acceptable" validity level for a simulation is a topic that deserves a separate discussion (see the end of this section).

Here, we discuss validity issues (which include the notion of fidelity described in Section 2), illustrate procedures for assessing a system's validity, and describe procedures for validating human judgment and decisionmaking models for use in simulations.

Validation Procedures

We categorize procedures for assessing a simulation's credibility into those that employ experimental techniques that satisfy scientific criteria of hypothesis testing and those that do not. Procedures in the latter category, however, do include the careful scrutiny of a simulation's modeling process for logical and physical errors. The first category of procedures we label *validation,* the second set *pretest,* which includes fidelity assessments. The testability criterion requires

- manipulating multiple factors in experimental designs that allow tests of causal hypotheses about independent and interactive effects on simulation outcomes (e.g., operational outcomes, decisions), and
- experimental designs to be matched in simulation and comparison systems.

Advantages to multiple-factor designs that are matched in simulation and comparison systems for assessing a simulation's validity are the ability to

- determine major independent and interactive causes underlying observed discrepancies between a simulation and selected comparison system, and
- pinpoint *location, magnitude,* and *direction* of discrepancies in outcomes between simulation and selected comparison systems.

The ability to locate the causal factors underlying discrepancies as well as to estimate their magnitude and direction aids the calibration process in achieving greater validity levels. In cases where calibration efforts fail, discrepancy information can be manipulated in controlled experimental designs to conduct *judgment* experiments that address the importance of obtained discrepancies given the intended use of the simulation. Procedures for conducting such judgment experiments are described at the end of this section.

Experimental design procedures presented here that meet the above testability criteria have broad application to the modeling and simulation community in general. They would also be useful to the test and evaluation community for assessing weapon system effectiveness.

Many researchers assessing the credibility of simulations depend almost exclusively on pretest procedures (discussed next) for their "validity" assessments with the idea that these procedures provide cause-and-effect kinds of information that they cannot in fact yield. We think it important to point out the difference in value of information obtained from the two categories of assessment procedures in terms of what is known about a simulation's credibility.

In this report, we reserve the term validation for the use of controlled experimental frameworks to test causal hypotheses—frameworks that satisfy scientific testability criteria.

Pretest procedures would generally be employed prior to and in preparation for hypothesis testing, and in fact would be expected to provide important information about what hypotheses might be more rigorously investigated. These procedures are described first before our discussions and illustration of the more rigorous validation framework.

Pretest Procedures

Pretest procedures for assessing a system's credibility can be characterized as nonexperimental, even though they include tests between point estimates (e.g., means, variances). When these procedures involve careful scrutiny of a simulation's modeling process for logical and physical errors, we refer to them as fidelity assessments.

Testimonials and "face validity." Testimonials are reports provided by experts or others. When they are credibility reports based on appearance, their status is referred to as "face validity." Reports could concern the gamut of internal or external validity issues outlined above. For example, in observing the unfolding of a simulated battle, an observer might report on the degree of realism of the behavior of the automated forces in battle scenarios or the degree of performance consistency of like automated and simulator (human-operated) weapon systems. A helicopter operator might report on the realism of a simulator's handling qualities, the realism of visual database features, or the adequacy of database cues for performing tactical maneuvers. A commander, upon being queried, might provide a researcher with information about his process and subsequently report on the accuracy of a set of decision rules that resulted from the query. Often, testimonials based on face validity are acted upon by programmers. Changes or adjustments are made in the simulation, observations are repeated, and readjustments are made in a feedback procedure until apparent "validity" has been achieved. Since no variables have been controlled, it is not possible to *test* the "correctness" or validity of the cause-and-effect implications contained in the reported observations or testimonials.

Statistical relationships among performance outcomes of two or more "equivalent" systems include correlations and tests between point estimates (e.g., means, variances). Point-estimate tests might focus on comparing mean performances (perhaps for a number of dimensions) of (a) an AVTB automated helicopter with that of its AVTB human-operated simulator counterpart for an internal validity check within the AVTB environment, (b) the CSRDF human-operated helicopter with its AVTB human-operated counterpart for a validity check between AVTB/CSRDF environments, or (c) a simulation system and its engineering specifications (prevalidity check). Or, correlations might be computed for two systems along specified dimensions. These statistical tools might be used to initially explore important fidelity questions that concern the physics of systems such as helicopter simulator flight characteristics (e.g., maximum speed, and bank and roll rates) or characteristics of weapon system components (e.g., trajectory, speed, and lethality envelopes of missiles, or contrast and resolution features of sensor systems). Since these statistical procedures do not feature simultaneous manipulation of multiple variables, it is not possible to uncover causal reasons behind observed discrepancies (or samenesses) in results, which makes it difficult to direct calibration efforts.

As mentioned above, it is expected that assessments performed in this pretest framework will generate many specific internal and external validity hypotheses that will require testing in a more rigorous framework.

Validity Tests for Simulation Outcomes

Figure 4.1 outlines the major ingredients in validating simulation outcomes. The schema consists of (a) the experimental design that specifies how the validity test will be conducted, (b) the simulations to be assessed (e.g., AVTB and CSRDF), and (c) the validity-base system that, by definition, constitutes the standard or "real world" against which to compare the simulation(s). Validity-base candidates might be any one or some combination of

- real-world system counterpart (shaded area in Figure 4.1),
- specifications,
- prototype,
- engineering model,
- live exercise or field test,
- small-scale model,
- contractor's simulator,
- combat simulation/human interactive model, and
- some other system.

A credibility argument has to be made for any system selected to serve as a validity base. Ideally, such an argument would be based on past validity *tests*. However, accessibility, cost, fidelity, or some combination of these reasons might also be factors in its selection. Once selected, that system is the standard against which the simulation under assessment is to be compared. The notion of using validity-base systems has been criticized because they are artificial in that they never consist of the full context of the real world. However, the real world can rarely suffice for a validity base because it is generally not possible to control variables in the real world without making it somehow "artificial." The irony is that the artificial world that allows the researcher to manipulate variables provides important information about what causes observed events in the real world. Thus, selection of the "best" artificial world (validity base) against which to pit the simulation being assessed is a necessary first step in validity tests; the validity base, experimental design, and analysis of experimental results determine the empirical validity of a simulation.

RANDMR340-4.1

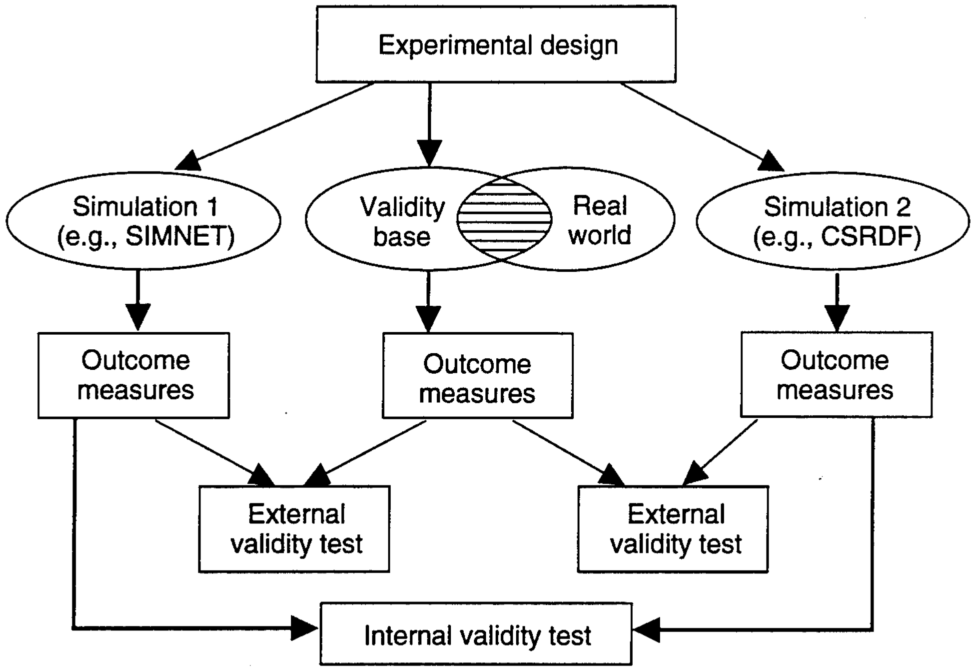

Figure 4.1—Validation Schema

Selection of the experimental design requires specifying the hypothesis to be tested, defining the outcome measures and multiple factors hypothesized to affect them, and specifying the simulator and validity-base trials needed to test the hypothesis. Trials specified by factorial combinations of variables allow identification of performance discrepancies between simulation systems or simulation and validity-base systems *when* the experimental design is matched in the systems being compared.

Different simulations such as AVTB and the CSRDF are expected to have variables that can be manipulated in one simulation and not the other because of the nature of the simulations' construction; for example, the level of "heat" emanating from objects in the environment can be manipulated in the CSRDF but not in AVTB, since the latter's networking system does not contain such a feature. The inclusion of such a variable in the experiment applied to that system may provide important calibration information. This is illustrated in the example presented below. When data demanded by the experimental design are collected and compared, conclusions can be drawn about the conditions under which the simulation is similar to its comparison system and where it requires changes for those variables manipulated in the experiment.

A Validity Test for Two Sensor Systems

For illustrative purposes, let us suppose we wish to test for validity in detection capability between the IR sensor in the CSRDF's advanced helicopter simulator and the IR sensor in an AVTB helicopter simulator. At the outset, it is known that AVTB does not account for radiance of weapon systems or other objects in the environment. The AVTB IR sensor presents a clear icon for an image, whereas the CSRDF IR sensor that accounts for heat patterns presents a "fuzzy" image. It may be the case that these differences in appearance produce unacceptable discrepancies in detection performances between the two systems. To exemplify how one might address the question of what difference in detection probabilities these projected images make, we selected an experimental design that manipulates three variables: **radiance** in the CSRDF, **number of nontarget vehicles** present in the target area in both the CSRDF and AVTB systems (a variable that should have a clutter effect for both the CSRDF and AVTB systems and an additional degrading effect for the CSRDF system in that more objects are radiating heat in the target area, thus possibly increasing the "fuzziness" of the CSRDF IR image),[2] and **threat intensity** in both the CSRDF and AVTB systems.

The idea in detection experiments is that some factors, for example, threat intensity, affect only an observer's *tendency* to report a target, whereas other factors—for example, number of nontarget vehicles in the target area and radiance—affect only an observer's ability to *perceive* the target. Both kinds of factors need to be included in a detection experiment to obtain an understanding of a sensor's capability (Green and Swets, 1966). The experimental design shown in Figure 4.2 exemplifies the techniques we have been discussing. In this design, the number of nontarget vehicles has three levels (high, medium, and low), threat intensity has three levels (low, medium, and high), and radiance has three levels (cool, warm, and hot). The selected levels of each of the three factors are fully crossed; that is, each level of each factor is combined with each level of every other factor, yielding a complete factorial design of 27 experimental situations. Data for the 2 × 2 matrix shown in Figure 4.2 would be collected for each of the 27 experimental trials for the CSRDF IR sensor; data for only nine trials (threat level (3) × number of nontarget vehicles [3]) would be collected for the AVTB IR sensor, since radiance is not a feature of that system. Data would be probabilities of correct detections and incorrect detections (false alarms [FA]—calling a nontarget a target) computed across detection reports obtained from a number of

[2]Radiance would be manipulated only in the CSRDF since, as mentioned, SIMNET is indifferent to differences in radiance levels.

RANDMR340-4.2

Threat Level	Radiance level: cool			Radiance level: warm			Radiance level: hot		
	Number of nontarget vehicles			Number of nontarget vehicles			Number of nontarget vehicles		
	High	Med	Low	High	Med	Low	High	Med	Low
Low 1									
Medium 2									
High 3									

Report	True: T	True: $\overline{T}$
"T"	p(detection)	p(FA)
"$\overline{T}$"	p(miss)	p(correct rejection)

Figure 4.2—Experimental Design for Consistency/Validity Test for Two Sensor Systems

operators for both the CSRDF and AVTB IR sensors. (The other two cells shown in the 2 × 2 matrix at the bottom of Figure 4.2—probabilities of misses and correct rejections—can be computed from this information.)

Hypothetical data that might result from fielding the experimental design in Figure 4.2 are presented graphically in Figure 4.3; probability of detection is plotted on the y-axis as a function of number of nontarget vehicles with a separate curve for each radiance level and a separate panel for each threat level. False alarm probabilities are shown by open squares in each panel. As can be seen, false alarms depend only on level of threat in this hypothetical example; they increase with increases in threat level, as do probabilities of detection (note that the position of the set of curves rises from panel A to B), supporting the hypothesis that threat level only affects an operators' *tendency* to report a target. Separations between the lower three curves represent the effect of radiance on target detection in the CSRDF system. Clearly, radiance has a significant effect on detection; it makes a larger difference when number of nontarget vehicles is high than when this number is low (compare the vertical distances at the high number of nontarget vehicles with that of the low number in each panel). The effect of radiance on detection decreases as threat intensity increases (separations between the curves decrease from Figures 4.3A to 4.3B).

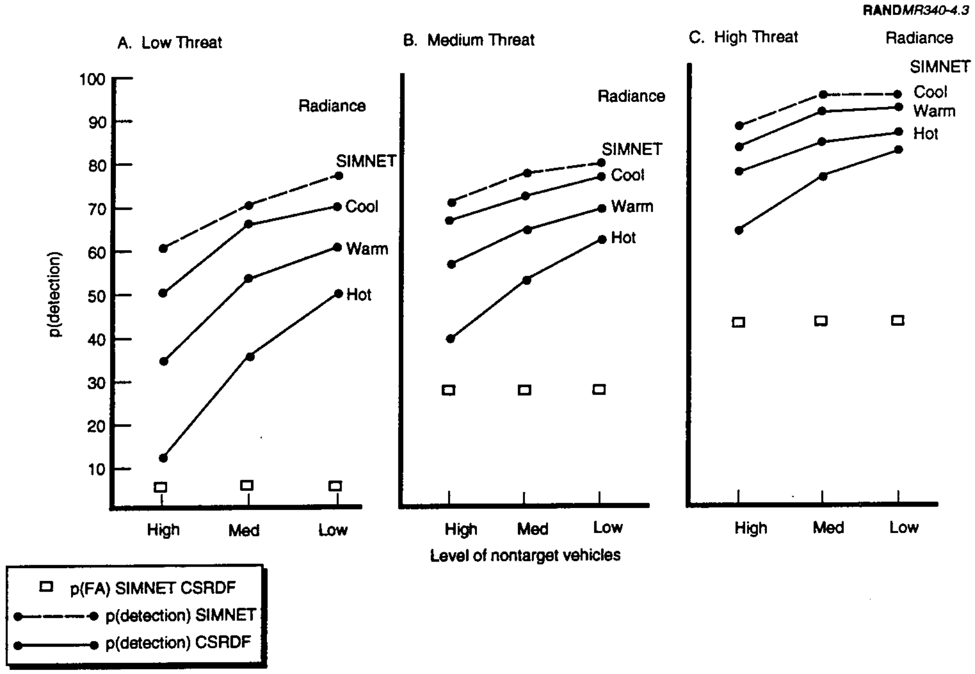

Figure 4.3—Hypothetical Data for Design Shown in Figure 4.2

The slopes of the hypothetical data curves in Figure 4.3 indicate the effect on detection probabilities of the number of nontarget vehicles in the target area. This factor affects probability of detection in both the AVTB and CSRDF systems. As threat level increases, this effect decreases (i.e., the difference in steepness of the four slopes decreases from Figures 4.3A to 4.3C). Under high threat conditions (Figure 4.3C), both the effects of radiance (separations between the curves) and number of nontarget vehicles (slopes of the curves) have significantly decreased and false alarm rates have increased, indicating that when operators are in high-threat conditions, the threat situation dominates the reporting; that is, increased threat causes operators' reporting tendency to increase.[3]

The data in Figure 4.3 can be replotted as receiver operator curves (ROC) so that tradeoffs between detections and false alarms as well as differences in discriminability between the CSRDF and AVTB IR sensors can be more directly

[3]An analysis-of-variance test applied to these data would confirm the statistical significance of the effects. As is the case with all statistical tests, however, the statistic does not reveal the nature of the independent or interactive effects; they are "blind" analyses. When data are plotted as in Figure 4.3, effects of factors can be seen. The ability to *observe* the direction and magnitude of effects is information that is crucial to validity assessments. Such observations are also critical for determining subjective measurement models (described in the next subsection) that allow the subjective importance of observed discrepancies to be measured.

assessed (Green and Swets, 1966). A subset of the data shown in Figure 4.3 is displayed as ROC curves in Figure 4.4. The subset consists of the AVTB data (dotted curves in Figures 4.3A and 4.3B) and the "cool" CSRDF radiance data (top solid curve in Figures 4.3A and 4.3B). In Figure 4.4, probability of a correct detection is plotted on the y-axis as a function of probability of a false alarm on the x-axis; dotted curves are the AVTB data, solid curves are data for the CSRDF under cool radiance levels. The lowest, middle, and highest solid and dotted curves are for low, medium, and high levels of number of nontarget vehicles. The separations between curves represent differences in operators' ability to discriminate targets under different numbers of nontarget vehicles; the points along each curve represent the difference in operators' tendency to report a target. The points increase both vertically (probability of detection increases) and horizontally (probability of false alarm increases) as threat level increases. In a detection experiment, it is possible for inconsistencies to be found in response tendencies (placement of points along the curve) as well as discrimination capabilities (separations between the curves). In this hypothetical example, inconsistencies between the two systems were found only in discriminability.

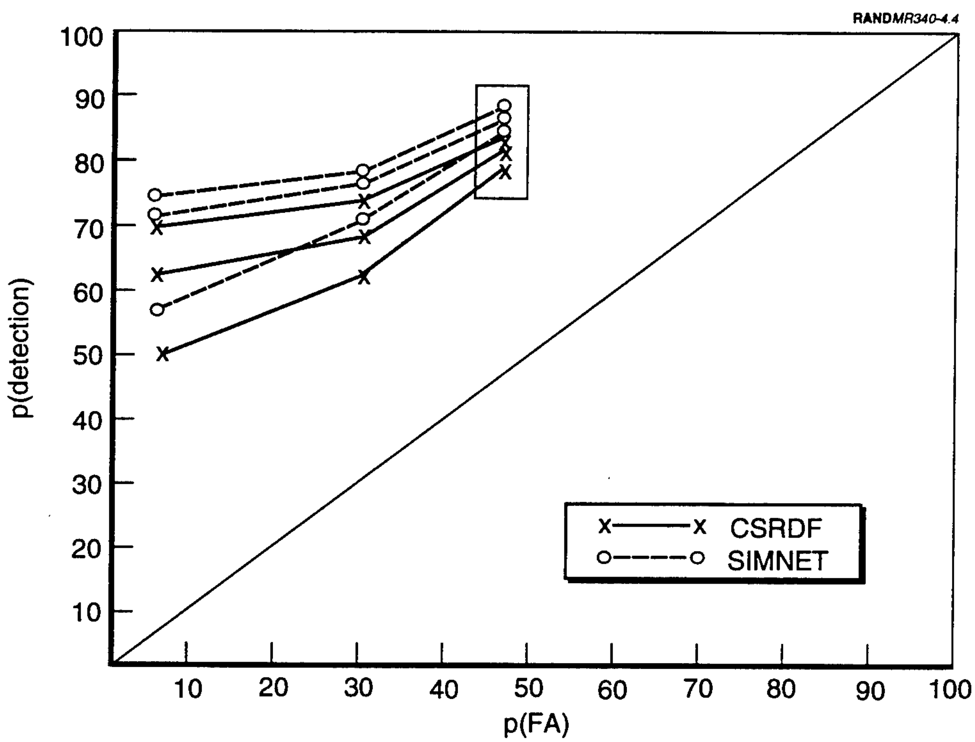

Figure 4.4—Receiver Operator Curves for Data Shown in Figure 4.3

Displaying data as ROC curves allows the analyst to see what situations were perceptually alike to the operators in terms of detecting targets. For example, the high level of number of nontarget vehicles in the CSRDF (top solid curve) produced almost identical probabilities of detection as the medium level of number of nontarget vehicles in AVTB (middle dotted curve). Other similarities in discriminability can be seen from these ROC curves.

Validity checks between the AVTB and CSRDF detection probabilities can be made from both kinds of data plots (Figures 4.3 and 4.4). Upon examination of these hypothetical data, it is clear that large discrepancies occurred under most conditions. However, the AVTB system performed very much like the CSRDF in a high-threat situation when the radiance level was cool (Figure 4.3C and boxed points in Figure 4.4). Detection probabilities obtained in AVTB differ from those obtained in the CSRDF under these conditions by only .02. The implication of these validity results is that, for situations exhibiting a cool background and high-threat environment (e.g., night mission analyses) useful analyses could be conducted without calibrating the two simulation IR sensor systems. For any other operational environment, however, it seems that calibrations would be necessary. Such calibrations might involve incorporating radiance and image "fuzziness" characteristics into the AVTB sensor system. Any such changes to the AVTB sensor would be limited by its design. If this limitation precluded achievement of acceptable internal validity levels, it would be necessary to alter the CSRDF's image (lower its fidelity). After changes are made to one or both systems, the experiment would be repeated to determine the effects of the calibrations. This validation/calibration process would continue until acceptable internal validity levels were obtained under all conditions specified by the experimental design or further calibrations were not possible, at which point the conditions under which the integrated product could be useful would have to be specified.[4]

Validity Checks

Before or during an internal validity experiment like the one just described, it would be wise to make an external validation experiment for one or both of the sensors. The external validity data could then help guide the direction of calibrations, thus preventing achievement of internal validity between the AVTB

[4]The reader is reminded that the IR sensor internal validity example illustrated in this section used hypothetical data and was constructed solely for illustrating the experimental design and graphic display analysis techniques discussed in this section.

and CSRDF sensors when the probabilities of detection and false alarms in the integrated product are inconsistent with a desired validity base.

An external validity test using the experimental design shown in Figure 4.2 might be conducted using a real IR sensor in a similar environment to that used in the CSRDF and AVTB simulators. Because of probable cost and feasibility constraints, one might have to use a subset of the experimental design shown in Figure 4.2. Validity and simulator data would be plotted on the same graph as shown in Figure 4.3 to show location as well as magnitude and direction of simulator/validity-base deviations. If just the CSRDF were used in the validity study (because it has greater fidelity and because it incorporates more IR variables), it would be calibrated to the validity-base sensor. The CSRDF IR sensor could then become the validity base for AVTB IR sensor calibrations that would result from the AVTB/CSRDF IR sensor internal validity tests. The important point here is that, when possible, external validity data should guide calibrations dictated by results of internal validity tests.

Comments

Given the hypothetical discrepancies in detection probabilities observed in Figure 4.3 between AVTB and CSRDF IR sensors, a higher-order question would ask what difference the observed differences make on a more global outcome such as number of targets killed or loss exchange ratios. To address such questions, one would want to use a few relevant variables in an experimental design like that illustrated in Figure 4.2.

The multiple-variable designs and graphic analyses presented here can also be profitably used to systematically assess how differences in the "same" sets of cues and scenarios in different databases produce differences in behavior such as the ability to perform various flying maneuvers or operational outcomes such as survivability, lethality, or kill ratio. Validity interests (speed, maneuverability, detections, engagements, and kills) between automated forces and manned simulator counterparts within the AVTB environment and between AVTB's and CSRDF's manned helicopters can be constructively explored in terms of causal factors underlying specified outcomes using the matching controlled experimental design framework illustrated here.

Validating Human Judgment Inputs to Simulations

A substantial proportion of battle simulations consists of human judgment input. Judgment input typically consists of decision algorithms or estimates of event occurrences incorporated in "look-up" tables. These human judgments rarely undergo validation scrutiny. Judgment data are usually obtained using procedures analogous to those described in the discussion on pretests. Thus, the "truth" or "validity" of interpretations given to the data is definitional (e.g., the "expert" says that is how he/she *thinks* about the problem). However, human judgments can undergo validity tests when appropriate experimental designs are used to generate the questions posed to respondents.

The ideas behind validation procedures for human judgments are similar to those described for assessing causal factors underlying simulation outcomes—the use of experimental designs that simultaneously manipulate multiple variables. The goal is to develop an experimental design that allows tests of hypotheses about (a) factors that affect judgments or decisions, and (b) the measurement model that underlies these effects. The emphasis is on constructing experimental designs that make it possible to detect and hence reject incorrect hypothesized measurement models. Algebraic modeling procedures developed in psychology over the last four decades provide a framework for testing among viable algebraic measurement models (Anderson, 1970, 1981; Birnbaum, 1974, 1990; Krantz and Tversky, 1971; Krantz et al., 1971; Veit, 1978). The algebraic modeling framework is referred to as a measurement framework because the model provides empirically based subjective measures of its causal factors and responses when it has passed rigorous validity tests imposed on it by the experimental design. These measurements (and hence model predictions) extend beyond the factor levels (specific information) used in the judgment experiment to a factor's full continuum (e.g., the full continuum of radiance levels if this factor has been used in a judgment experiment).

Adequate tests among various algebraic models require different experimental design features that depend on the *unique* predictions of the models under investigation. The idea is to *reject* incorrect algebraic models (those that fail their validity tests) and use algebraic models that pass their validity tests to better understand how a particular situation affects a person's judgments and, thus, under what conditions those judgments might be expected to change. In addition to the references cited, examples of experimental designs that allow tests among hypothesized algebraic judgment or decisionmaking models can be found in Veit and Callero (1981), and Veit et al. (1984), as well as illustrations of a method (the Subjective Transfer Function Approach) that extends the modern measurement testability principles to the construction of measurement models of complex systems where many activities and processes occur simultaneously and in tandem (e.g., military command and control).

The notion of validating a judgment or decisionmaking model can be demonstrated with the following generic example. Suppose a military expert was asked to estimate the decision he/she would make in 20 situations that were constructed from a 5 × 4 (X × Y) factorial design like that shown in Figure 4.5. Suppose further that it was hypothesized that these experts combined the two pieces of information describing each situation (e.g., a1, b1 would be the two pieces of information included in the situation generated by the upper left-hand cell of the experimental matrix shown in Figure 4.5) according to a simple adding model, that is, that the whole (the response) was the sum of the separate parts.

RANDMR340-4.5

Y factor

X factor	Y_1	Y_2	Y_3	Y_4
X_1	X_1Y_1	X_1Y_2	X_1Y_3	X_1Y_4
X_2				
X_3				
X_4				
X_5	X_5Y_1			X_5Y_4

Figure 4.5—A 5 × 4 (X × Y) Factorial Design That Generates 20 Situations for Judgment

This model can be written as

$$D_{(ij)} = a\left[x_{(i)} + y_{(j)}\right] + b$$

where $D_{(ij)}$ is the decision provided by the military expert to the situation described by the *ith* and *jth* piece of information, x and y are the *subjective* values associated with the *ith* and *jth* piece of information, respectively, and a and b are linear constants, indicating that the decisions provided by the military experts are assumed to be a linear function of their subjective decisions. With this latter assumption, it is possible to *test* the additive hypothesis shown above by simply graphing the decisionmaking data obtained from the military experts. This graphic analysis, shown in Figure 4.6A, depicts what the decisionmaking data *should* look like if the additive model is correct; vertical separations between any two curves should be the same, independent of the value on the x-axis (the curves need not be linear, just parallel). If the data looked like those shown in Figure 4.6A (an analysis of variance test of the interaction between x and y should be nonsignificant), it would be concluded that the additive decisionmaking theory appropriately represented the military experts' decisions. The subjective values or measures associated with the X and Y factor levels would be derived from the theory (they would be the least-squares estimates under the additive model). The additive theory would provide subjective measures along the entire X and Y factor continua if factors defined physical dimensions (e.g., distance or time).

RANDMR340-4.6

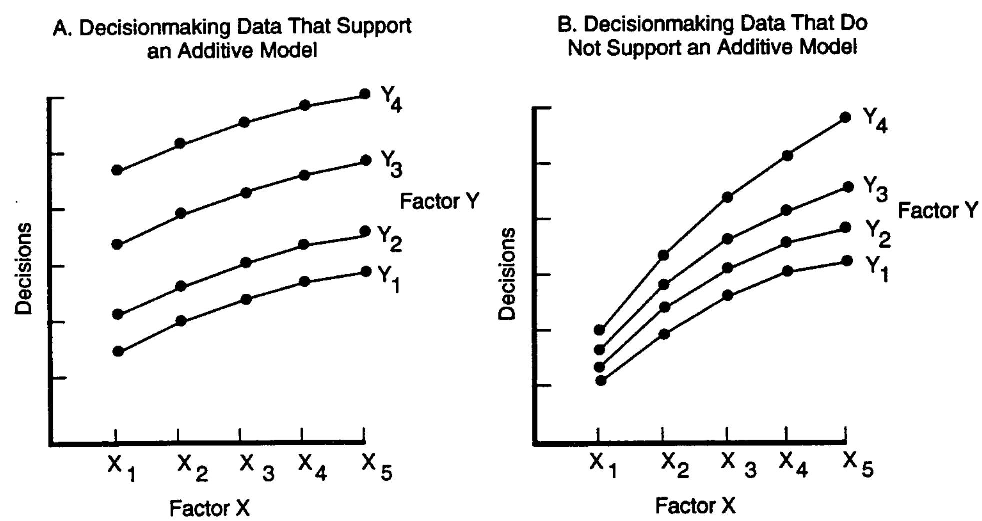

Figure 4.6—Graphic Analyses of Hypothetical Data (Decisions)

If the decisionmaking data graphed as in Figure 4.6B, it would be necessary to reject the additive model since these curves are not parallel but form a fan of curves that diverge to the right. When the decisionmaking theory is rejected, the subjective measures it provides of the factor levels and decisions are also rejected. Thus, the ability to measure effects of information on judgments depends on the ability to correctly model the judgments or decisions.

In practice, a researcher would want to field an experimental design that allowed a "correct" model to be determined, that is, allowed tests among many alternative decisionmaking or judgment theories. (In a validity test of a judgment or decisionmaking model, the judgment data serve as the validity base against which to test among hypothesized measurement models.) The experimental design shown in Figure 4.5 does not have the power to distinguish among many models. Experimental designs that provide more flexibility in determining appropriate models are illustrated in Anderson (1981), Birnbaum (1990), Birnbaum and Veit (1974a, b) and Veit et al. (1984), as well as some of the other articles on subjective measurement cited in this section.

A somewhat simplified description of the steps involved in testing an algebraic model's validity are

- hypothesize an algebraic model (or a number of models),
- construct an experimental design that permits tests among their unique predictions,
- define the respondent group to be modeled,
- collect judgment data (validity-base data) from the respondents according to the dictates of the experimental design,
- assess discrepancies between the structure obtained in the validity-base judgment data and those predicted by the hypothesized models using graphic analyses (see Figure 4.6), and
- reject models (and their measures) that do not predict the data structure found in the validity-base judgment data.

The ability to reject incorrect models lends credibility to the model that passes its tests and hence to that model's measures. Judgment and decisionmaking models serve at least four important functions in computer modeling and simulation. They can

- address research questions not possible in the simulation environment (e.g., combat effectiveness of advanced weapon system capabilities),

- provide information about the relative importance of observed simulation/validity-base data discrepancies such as those observed in Figure 4.3,
- answer questions about what factors affect users' judgments or decisions, and whether the factors affect the user the same in simulator and validity-base environments, and
- be parsimoniously embedded in computer simulations to provide subjective input data such as commanders' decisions on force deployment, engagement, and the like.

Often a validity-base environment cannot be cost-effectively created to answer validity questions because of difficulty creating projected weapon system or other capabilities, appropriate combat environments, etc. In such cases, information about outcomes can be obtained by systematically examining the opinions of experts so that their decisionmaking and judgment processes can be modeled. In the next few paragraphs, we briefly expand on some of the uses for judgment and decisionmaking models in computer modeling and simulation.

Decision or Judgment Models

When simulations contain decision algorithms or tables, it is especially important that these come from models that have received empirical support for their validity. Once an algebraic decision model has received such support, it can be incorporated into the simulation, thus increasing the credibility of the simulation's outcomes (Veit, 1994). Further, validity between simulations that require the same decision rules is simplified by incorporating the same decision model into both simulations. Algebraic models have an advantage in addition to being testable when appropriate experimental designs are used. They can be translated into a simulation's language (and thus incorporated into a simulation) using very few lines of code.

Perceptions of SIMNET/CSRDF Database Objects

Perceptual models can be developed to provide information about the similarity of the "same" objects in different databases in terms of, for example, their distance and size from different observation points, or the location in time and space of both fixed and moving objects. Perceptual models of database phenomena developed in different databases (CSRDF, AVTB, a validity base) can be matched for validity, using graphic analyses like those shown in Figure 4.3.

Assessing the Importance of Observed Inconsistencies

What is a reasonable and credible criterion for deciding on the sufficiency of achieved internal or external validity levels or the importance of observed discrepancies? As mentioned earlier, operators who perform simulation tasks in the real world can make these judgments. When "experts'" judgments are obtained in a structured experimental framework that permits tests among judgment models (and thus, rejection of incorrect models), an analyst is provided with information about the causal underpinnings of the "experts'" opinions—the reasons (associated with the simulation) underlying the judgments and, hence, the credibility of the simulation.

Concluding Remarks

Section 4 has discussed calibration combined with both internal and external validity issues in computer modeling and simulation. The term "validity" is widely used in the computer modeling and simulation community to refer to the credibility of a simulation. It connotes real-world "truth," and thus seems an inappropriate term for characterizing the potential status of any hypothesized state of the world of which a simulation is one. We view validation procedures as procedures for assessing the credibility of a hypothesis; assessment procedures that provide the most credible cause-and-effect kinds of information about observed events are those that employ controlled research paradigms. Data from such paradigms can provide evidence for or against a hypothesis; they cannot *prove* a hypothesized state of the world to be correct.

Factorial designs, matched in simulation and comparison systems, were proposed as the basic experimental design for testing cause-and-effect hypotheses about those systems. As the number of factors and factor levels increases in factorial designs, the number of experimental trials increases. This is a feature of factorial designs that has been used to argue against their use. However, their advantage is both the amount and the quality of information obtained about the situation under investigation. Experimental designs for assessing both internal and external validity, such as the one illustrated in Figure 4.3, should be unconstrained by practical considerations (e.g., time and resources) that may be imposed on a particular research effort. Constructing "ideal" experimental designs is an important initial step in any data-gathering situation because the experimental design serves as a guide to selecting the data-gathering trials that would provide the most valuable information about the questions at issue under the constraints of the particular data-gathering effort.

5. Proof-of-Principle Demonstration

Significant benefits could accrue to the weapon system development and acquisition process from integrating high-fidelity R&D simulators of the systems of interest into a simulation network that provides large-scale operational environments featuring the human element in battlefield interactions and command and control decisions. Such an integrated system would have the potential to support the entire process and could greatly increase its quality in terms of analytic validity levels, and hence the usefulness of the analyses and operational test and evaluations. With such potentials it is appropriate that a proof-of-principle demonstration be undertaken to determine if, indeed, such an integration could be accomplished within reasonable time and cost constraints, and the final integrated system could provide the quality of analytical support anticipated.

We recommend that a full proof-of-principle demonstration be undertaken to integrate the CSRDF with the AVTB with the following objectives:

1. Establish the feasibility of integrating a high-fidelity R&D simulator with a SIMNET-based DIS environment capable of supporting credible weapon system analysis within reasonable time and cost constraints.
2. Develop integration tools and procedures that will transfer to future similar integrations.
3. Upgrade internal and external validity properties of selected AVTB/SIMNET representations to meet weapon system analysis requirements.
4. Demonstrate the breadth and flexibility of a DIS-based analysis.
5. Identify additional modeling and simulation validity shortfalls for future upgrade.

Establishing the feasibility of integrating these disparately developed systems in such a way that the resulting system is significantly better than existing analytical tools could open the door for a new era of weapon system analysis. It should be noted that, while the integration itself poses the most difficult technological problems, accomplishing this by producing an integration "tool kit" (e.g., generalized database transfer programs) and a high-quality analysis capability are the more fundamental purposes.

Throughout the proof-of-principle integration, the integration should proceed in such a way as to develop tools and procedures that would form a basis for integrating disparate simulators with SIMNET-based systems in particular, and with other DISs in general. Potential areas where broadly applicable tools and procedures could be developed include database interchange, consistency and validity evaluation, protocol formulation, and gateway architecture. If development of these tools and procedures is successful, future integrations should require significantly less preparatory intellectualizing and development of supporting integration software and hardware systems. A "how to" manual would be a useful product from this effort.

AVTB and other SIMNET-based systems suffer from fidelity and internal validity shortcomings that should be corrected before being used for analysis. The most critical shortcomings would have to be corrected in the proof-of-principle demonstration to achieve the enhanced analytical capability desired. This objective may be satisfied even if great difficulty is encountered in meeting all the CSRDF integration and high-level analytical capability objectives. Future users of AVTB and other SIMNET-based systems would benefit from this effort.

An area needing priority attention in the SIMNET-based systems is the modeling of phenomena such as electronic warfare, dynamic signature generation, dynamic (as opposed to fixed) terrain databases, atmospheric conditions, day/night effects, and many behavioral representations. The difficulty of adequately modeling these phenomena lies in the inability both to select or develop credible models to incorporate in the systems and to obtain believable empirical data to use in models. An identification effort should specify the highest-payoff areas for improvement and methods for accomplishing the necessary modeling upgrades. Current activities within the distributed interactive modeling community that address these and similar issues should be considered and capitalized on to the extent possible.

Finally, the resultant product would also provide an enhanced laboratory to conduct rotorcraft research and development that would prove valuable to the defense community (particularly the Army and ARPA) and NASA.

Connectivity

A number of weapon systems have achieved connectivity with DISs in the sense that intercommunication could take place. (This limited connectivity is to be distinguished from the integration goals of the proof-of-principle recommended here.) Notable among these is the preliminary connectivity established between the CSRDF and AVTB in March 1993 and again in May 1993 to showcase distributed interactive simulation potentialities.[1]

The CSRDF program architecture was modified for asynchronous processing and data interchange with network elements; communication interface units were developed and a replica of the CSRDF fixed terrain base was manually installed on AVTB. Although this concept demonstration activity achieved its demonstration purposes, it did so without addressing the critical technical issues associated with weapon system analysis (e.g., fidelity, consistency, correlation, validation), or providing a tool kit for future integrations, the two goals of the proof-of-principle described here. Nonetheless, the CSRDF/AVTB interface developed for demonstration purposes provides the starting point for the proof-of-principle integration. It includes the communication and computational hardware and software that permits the CSRDF to reside as a node on the AVTB network, transmit protocol data units (PDUs) that inform the other nodes as to the CSRDF's dynamic state, and receive and interpret PDUs from other nodes to determine those nodes' characteristics (e.g., location, movement, activities such as firing or being engaged) and present appropriate sensor and visual displays to the CSRDF crew. This connectivity will likely need to be enhanced to meet the extended requirements of a weapon system analysis. Hence, determining these enhancements and implementing them will be part of the proof-of-principle effort.

Database Interchange

In conjunction with and as an essential part of the proof of principle would be the development of fixed databases and scene generation (terrain, foliage, and cultural features) that are consistent between AVTB and CSRDF. For analysis of aircraft directly interacting with ground forces, AVTB would be the central system that a variety of developing weapon system simulators would use to analyze their operational performance. Hence, the concept of AVTB as the source of fixed databases in any integrated system should be adopted. Within that concept, a standard database interchange for transferring AVTB (or other SIMNET-based systems) fixed databases to other systems should be preselected (for examples, see STRICOM's Summary Report, Vol. 1, 1993). The integration efforts associated with this issue would focus on achieving the database

---

[1]The May 1993 connectivity supported the Association of the United States Army (AUSA) Louisiana Maneuvers and Army Simulation Initiatives Symposium and Exhibition at Orlando, Florida (24–26 May 1993).

interchange itself and developing consistent image-generation overload management algorithms.

Database Interchange Activity

Databases for integrating the CSRDF CompuScene IV with the AVTB network would need to be developed in several steps. An acceptable sample subset of a Hunter-Liggett database could serve for the integration effort. This database would be encoded in the interchange format and provided along with a software library for encoding and decoding the dataset. Next, software would be developed to convert the dataset into a format usable by the GE modeling tools to build out-the-windscreen (OTW) and thermal databases for CompuScene IV. Once the conversion issues were resolved, the augmented Hunter-Liggett database that includes colors and textures could be encoded in the interchange format for conversion to GE tool formats.

Finally, building the OTW and thermal databases for the CompuScene IV will probably entail some divergence from the reference model provided in the interchange format. The key issue relating to interoperability between the high- and low-resolution simulators will involve questions of intervisibility. An "intervisibility equivalence" measure must be defined that determines when the GE and the AVTB databases agree on intervisibility. This measure could be used, if necessary, to guide further refinement of the databases to achieve a satisfactory level of interoperability, determined by using appropriate consistency evaluation techniques.

Coordination of CIG Overload Management Algorithms

CIG systems inherently must live within a stringent processing limit associated with their designed characteristics. Careful management of loading (pixels and polygons) is often employed to maintain optimum loading in CIG processing. Distributed simulation plays havoc with the visual loading calculations for visual systems. For example, a given display may suddenly have to deal with hundreds of moving models and ballistic effects because of the actions of other manned and unmanned simulators. Any given CIG system will face overload conditions at some time. These overloads must be coordinated appropriately for different CIGs, or closely correlated CSRDF and AVTB weapon systems (e.g., helicopter teams) could have different views of the battlefield.

The SIMNET database was designed to reserve approximately 50 percent of its processing power for dynamic objects such as vehicles and special effects such as tracers and bomb bursts. The processing of these dynamic objects determines when the SIMNET CIG will overload. Currently, the SIMNET manned vehicle simulators have been programmed with priority tables that determine the order in which vehicles and effects should be dropped if overload appears imminent. In addition, the SIMNET CIG uses a "graceful overload" image degradation method, processing objects and terrain from front to back. The CIG currently has no direct way of sensing polygonal or pixel overload, so it is possible for important database elements to disappear without feedback to the simulation or CIG real-time software.

The higher-fidelity CSRDF CIG system is currently restricted to processing 31 dynamic objects, including vehicles and special effects. Even though this capability may be increased in the future, problems of integrating a system that processes fewer dynamic objects with the lower-fidelity SIMNET system remains the same. One way to rectify disparate processing levels is to increase the capability for the more limited CSRDF system through software modifications. If this is not possible, the SIMNET CIG may have to be restricted to processing a similar number of objects for the proof-of-principle demonstration. In any case, it will be necessary to reach agreement regarding which objects, terrain levels of detail, and special effects are highest in priority. This will require careful database planning, engineering estimates of processing available for dynamic objects, and coordination of the prioritized filter.

AVTB Fidelity Enhancements

As repeatedly noted in this report, SIMNET-based systems suffer from fidelity deficiencies that seriously limit their usefulness in conducting operational analyses. The proof-of-principle demonstration could not feasibly correct all such fidelity deficiencies but would focus on those that directly affect the selected analytical problem. In general, deficiencies that would most likely affect helicopter capability analysis and thus need resolution would be sensor modeling and visual image representation and semi-automated force modeling.

Sensor and Visual Images

The modern battlefield will be dominated by sensors, and credible operational analysis will depend in large part on the credibility and fidelity of the portrayal of sensor capabilities. Section 3 discussed sensor modeling and visual image representation technical issues in considerable detail. For weapon system effectiveness analyses, the importance of these issues resides in how they affect operational performance or, in an integrated system, how they skew the apparent capability of disparate simulators such as AVTB and CSRDF helicopter simulations.

The AVTB representation of sensors and signatures on the battlefield is seriously flawed. For example, manned simulator sensors are restricted to a 7 km range whereas semi-automated force (SAFOR) sensors have unlimited range; their detection is constrained only by line-of-sight and entries in range tables (discussed below). Helicopter manned simulator IR sensor models do not account for IR clutter, false targets, indistinct imaging, target aspect or signature, or a number of other factors that affect IR detection. IR sensing deficiencies begin with vehicle signatures. On the IR display, the terrain appears with a green hue, roads and tree trunks appear light green, and a vehicle appears white or at least distinctly lighter. Even at extreme ranges, a vehicle will spotlight clearly. By narrowing the field of view to "zoom in," the operator sees a perfect icon, although the icon decreases in size as the range increases. Such an approach to modeling the IR sensor greatly overstates the performance of any actual sensor and the ability of an operator to detect and track targets. In contrast, the CSRDF GE CompuScene IV image generator models the scene, signatures, and sensor in great detail and uses IR transfer functions to present a detection problem for the operator more akin to using actual fielded sensors.

Guided by the validation and calibration procedures discussed in Section 4, both the direct effect of the limited AVTB sensor fidelity and differences in AVTB and CSRDF operational performance resulting from this capability disparity would have to be carefully assessed and resolved.

Semi-Automated Forces

Semi-automated simulations of friendly and enemy force elements and weapon systems are critical to the simulation of large-scale battlefield operations, since it will rarely be possible or even desirable to assemble the large numbers of simulators and crews needed to assess battlefield operations, particularly when repeated runs are envisioned to explore variations in systems' features and capabilities. To avoid problems of availability, efficiency, repeatability, and human learning effects, SAFOR can readily provide forces tailored to the analytical problem and, at least in principle, be programmed to execute well-defined behaviors repeatedly without tiring or learning.

The AVTB semi-automated force system capability provides for an operator at a workstation to create enemy and friendly units and specify operational criteria causing automated activity by forces that appear on the battlefield with precisely the same icons and dynamic effects as manned system simulators. Once set, SAFOR actions follow from algorithms that react to preset rules and specific parameters input by the commander.

Deficiencies of the current AVTB SAFOR center around simplifications in representing systems and phenomena that have led to grossly inadequate and erroneous modeling of sensor detections and system behaviors. The source of many problems is that one table specifies a single set of detection probabilities for *all* SAFOR ground vehicles (enemy and friendly) sensing *any* other vehicle (enemy or friendly). Similarily, another table specifies detection probability for all SAFOR air vehicles. Table entries are organized by range, whether the sensed vehicle is moving, stationary (fully exposed), or hull down, and by two sensor fields of view. Hence, for example, at any given range, a SAFOR tank has a single probablity entry for detecting a flying helicopter, a moving tank, or a moving bridging vehicle; and that probability applies to *every* ground vehicle with *any* type sensor or sensor suite on board. No distinction is made for the type or number of sensors mounted, the signature of the target, or battlefield phenomena such as obscurants. When a weapon system fires, the probability of detection is increased by a factor of unknown origin or validity.

This problem has now been recognized and improvements to semi-automated forces are being developed. One such improvement maintains the original concept but increases the number of tables from one to four in each of the ground and air tables to provide flexibility to reflect multiple, different onboard sensors but not to differentiate for target types or conditions. A major enhancement called MODSAF, under development by Loral, provides the framework for the user to implement representations of the user's choice by designing and programming modules that operate within the framework. MODSAF reportedly will be installed in selected SIMNET-based systems in the near future.

Guided by the validation and calibration procedures discussed in Section 4, the effect on credible operational performance analysis resulting from the particular SAFOR system in AVTB when the proof of principle begins would have to be carefully assessed and resolved.

Proof-of-Principle Activity Sequence

Figure 5.1 shows the activity flow to accomplish the tasks described. Although cross involvement by AVTB and CSRDF expert personnel would take place throughout the process, in the early phases most of the work would be independent. In the later phases of the integration, internal and external validity

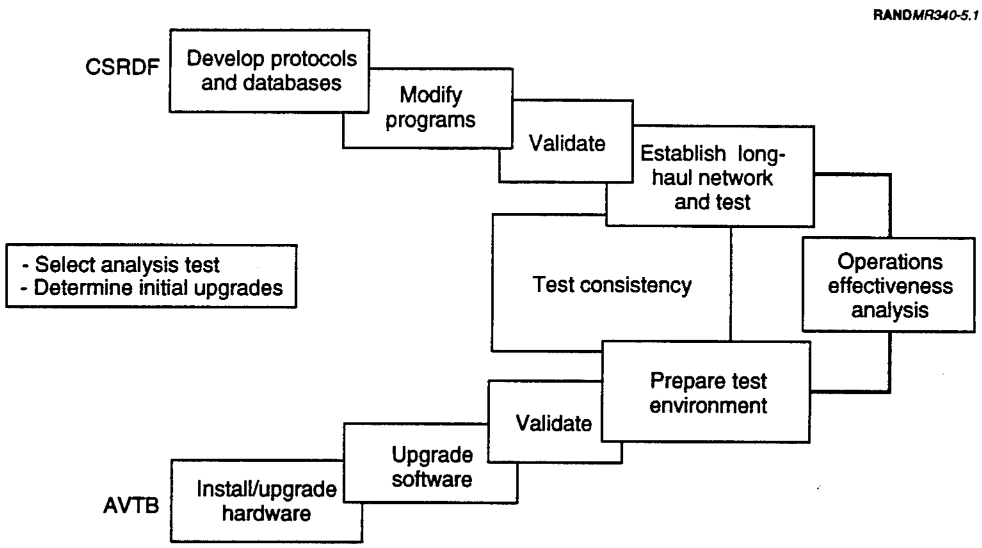

Figure 5.1—Proof-of-Principle Activities

evaluations would cut across both systems, as would the preparation of the systems to conduct the selected operations effectiveness analysis.

To design, prepare, and conduct the demonstration weapon system analysis would require several tasks. A suitable problem would be selected consistent with the capability of the integrated system—in essence, an advanced helicopter system or subsystem analysis. Tactics and operational procedures would be developed initially on the CSRDF local area network and finalized with the AVTB interface. The combat environment, including all friendly and threat systems, would be established in AVTB. For efficiency without loss of generality, the Hunter-Liggett test center fixed database now in both the AVTB and CSRDF should be used, smoke and dust cloud requirements should be avoided, and the entire analysis should be conducted in an unclassified mode.

Finally, like systems relevant to the analysis (e.g., an advanced helicopter or sensor suite) in CSRDF and AVTB and internal to AVTB would be calibrated (guided by validity data). Once the analytical simulations were complete, both AVTB and CSRDF would assess the results.

6. Conclusions

The complexity of modern warfare and the increased technical performance of modern weapon systems make it extremely difficult to adequately estimate with existing simulation techniques the capabilities of new weapon systems on the battlefield. Declining defense budgets accentuate the importance of making wise decisions on alternative weapon system choices or upgrades to existing systems early in the R&D cycle. Analysts need better ways of testing weapon system concepts and examining performance tradeoffs at a reasonable cost. Current analytic techniques used in the early phases of weapon system development—such as paper designs, engineering calculations, and combat models—are limited in fidelity and the real-world interaction of the human operator with hardware in a combat environment.

We have identified in this research what we believe to be a feasible approach to developing a simulation environment for weapon system analysis that can overcome crucial limitations in current analytic techniques. The approach involves integrating a high-performance weapon system simulator built for R&D into the battlefield simulation environment provided by an upgraded version of SIMNET. We believe that, following necessary internal and external validity work, this combination will provide the DoD with a greatly improved tool for analyzing the operational effectiveness of weapon systems. It should also prove quite effective in later phases of the acquisition process—prototyping and live testing—when choices can be made based on simulated hardware before prototypes are built, and operational tests can be designed without environmental and safety constraints, making them more relevant to combat operations.

We recommend that a comprehensive proof-of-principle demonstration, as outlined in Section 5 of this report, be undertaken so that this simulation approach can be made available to the defense research, development, and acquisition community. We believe this is a unique opportunity for advancing the state-of-the-art of weapon system analysis and, if successfully accomplished and used in the weapon system development and acquisition process, will greatly improve the DoD's ability to assess future weapon system capabilities.

Appendix

A. The Weapon System Development and Acquisition Process

The weapon system development and acquisition process for major defense acquisition programs[1] is divided into as many as five phases. Each phase is preceded by a decision point, referred to as a milestone, at which a Defense Acquisition Board (DAB)[2] and the Secretary of Defense (SECDEF) determine if and how the program should proceed. The phases are tailored to the specific program to minimize acquisition time and life-cycle costs and are consistent with the urgency of need, degree of technical risk involved, and progress as demonstrated by test results.

A program moves through each phase at its own pace. Prior to each decision point, there is a thorough review of program status by the relevant DAB committee[3] and staff briefings to various agencies within DoD. The DAB committee(s) makes a recommendation to the full DAB, which reviews it and other relevant information to make a recommendation to SECDEF. The SECDEF's decision is documented in an Acquisition Decision Memorandum (ADM), which includes approval of goals and thresholds. The ADM is issued by the DAE for execution through the Component Head.

The process is based on a continuous analysis of mission areas to identify deficiencies and opportunities and determine more effective means of performing tasks. The mission need is reviewed and approved at each milestone. There appear to be four primary acquisition policy objectives: (1) affordability (cost, priority, and resources); (2) program stability enhancement through realistic planning and estimating and minimization of changes to an ongoing program; (3) tailoring of the acquisition strategy to minimize time, including cost, schedule, and performance tradeoff studies, technological risk assessment, and

[1]Major defense acquisition programs are programs estimated to cost more than $200 million in Research, Development, Test, and Evaluation (RDT&E) and/or $1 billion in procurement (1980 dollars) or are specifically designated major by the Secretary of Defense.

[2]The DAB is the senior Department of Defense (DoD) review board and is chaired by the Under Secretary of Defense (Acquisition) in the role of the Defense Acquisition Executive (DAE). The vice chairman of the Joint Chiefs of Staff serves as the vice chair. The DAB assists the DAE with milestone reviews and policy formulation.

[3]The DAB is supported by ten committees organized along mission areas: Science and Technology; Nuclear Weapons; Strategic Systems; Conventional Systems; Command, Control, and Communications; Test and Evaluation; Production and Logistics; Installation Support and Construction; International Programs; and Policy and Initiatives.

competitive prototyping; and (4) competitive prototyping of critical components, subsystems, or systems and early testing and evaluation beginning in the demonstration and evaluation phase.

A description of the milestones and the phases' objectives and activities follows.[4]

Milestone 0: Concept Exploration and Definition

Milestone 0 decision: Approval/disapproval of a mission need and entry into concept exploration/definition phase.

Issues considered: Mission needs clearly presented as worthy of resource commitment; threat validation; needs expressed as performance requirements (not system description); affordability; constraints identified; acquisition strategy discussed; operational utility assessment; ability of modification or upgrade to perform mission rather than a new start.

Phase 0 objectives: Identify deficiencies in current capabilities and/or determine most effective means of performing assigned tasks; establish system performance requirements.

Phase 0 activities: Explore system design concepts; develop initial operational concept; tradeoff studies; risk analysis/assessment.

Milestone I: Demonstration and Validation

Milestone I decision: Approval/disapproval to proceed into concept demonstration/validation phase.

Issues considered: Threat validation; requirement validation; technical risks identified; criteria established for phase; acquisition strategy tailored to risks, etc.; thresholds and objectives established for next milestone; program alternative tradeoffs; cost, schedule, performance tradeoffs.

Phase I objectives: Select the preferred system concept; demonstrate that risk areas are resolved; demonstrate that only engineering (rather than experimental) efforts remain; demonstrate that resources are available.

[4]The process as described here is highly generalized and largely theoretical. In practice, it is tailored to reflect the characteristics of each weapon system program, implying that details of the process usually vary across programs. For instance, programs commonly skip or combine some phases. Further, the process described reflects recommendations of the Packard Commission that have been implemented to varying degrees by OSD and the Services.

Phase I activities: Prototyping; competitive demonstration in operational environment; finalize operational concept; review test and evaluation plan to identify areas of risk reduction in Operational Test and Evaluation (OT&E) planning; translate functional requirements to technical specifications.

Phase II: Engineering and Manufacturing Development

Milestone II decision: Approval/disapproval to proceed into engineering and manufacturing development phase, and as appropriate, low rate of initial production.

Issues considered: Technical solution described in detail; resolution of risk confirmed through testing; cost effectiveness analysis results; rejection of alternatives supported; future cost and schedule defined in detail; acquisition strategy defined in detail; program cost and risk vs. military value; development-production transition planning; program stability factors considered; results of prototyping and demonstration/validation phase; establish detailed cost, schedule, and performance goals/thresholds.

Phase II objectives: Complete system development; demonstrate that all technical, operational, and resource requirement thresholds are met; obtain independent OT&E assessment.

Phase II activities: Develop and produce limited quantity; development test and evaluation; independent OT&E; production plan in detail.

Phase III: Production

Milestone III decision: Approval/disapproval to proceed into full rate production and deployment phase.

Issues considered: All thresholds met; operational effectiveness confirmed; producibility risks reduced to acceptable levels; production quantity validated; production cost and affordability; reliability, maintainability, and logistics support.

Phase III objectives: Produce specified quantity at economical rates; deploy systems to operational units.

Phase III activities: Production; deployment; follow-on OT&E.

Milestone IV: Major Modification

Milestone IV decision: A review after initial deployment to evaluate operational effectiveness and assess the need for an upgrade or replacement.

Issues considered: Capability to meet original or evolved mission requirements; potential necessity of modification or upgrade to ensure that mission requirements are met and that the useful life is extended; changes in threat; changes in technology.

Phase IV objectives: Determine effectiveness of various alternatives to meet threat, including modification or upgrade to system or system replacement.

SOURCES: DoD Directive 5000.1, *Defense Acquisition Programs,* February 1991.

DoD Instruction 5000.2, *Defense Acquisition Management Policies and Procedures,* February 1991.

B. SIMNET

The technology underlying simulator networking, developed over the last decade, has been implemented as the SIMNET system. SIMNET was sponsored by the Defense Advanced Research Projects Agency (DARPA) in partnership with the U.S. Army. The program's goal was to develop the technology for networking large numbers of interactive combat vehicle simulators and their supporting elements. In effect, SIMNET provides a simulated world in which fully manned platoon-, company-, and battalion-level units can fight force-on-force engagements against an opposing unit of similar composition. SIMNET provides a joint, combined arms environment with the complete range of command and control and combat service support elements essential to actual military operations. All of the elements that can affect the outcome of a battle are represented, and victory is likely to go to the unit that is better able to plan, orchestrate, and execute its tactical operations.

SIMNET terrain is a somewhat simplified representation of actual terrain, constructed from Defense Mapping Agency data. It is represented with sufficient realism that the crews can navigate through it, recognizing roads, rivers, hills, tree lines, and other distinctive terrain features as they would on actual terrain. In addition, they see combat vehicles, combat support vehicles, and combat service support vehicles—tanks, helicopters, self-propelled howitzers, fuel trucks, etc. The actions of these vehicles reflect the control actions of other vehicle crews in other simulators elsewhere on the SIMNET network.

The simulators on the SIMNET network may be located on the same local area network (LAN) or on other LANs that are linked in a wide area network (WAN) by a variety of techniques. Network traffic measurements indicate that the SIMNET network can support up to 1000 active vehicle simulators on each LAN.

Distributed Simulation Architecture

The concept of distributed simulation is central to the SIMNET architecture. In SIMNET, there is no central computer directing the activities of the various simulation elements. Each simulator has its own microcomputer, which is in continuous communication with each of the other simulation elements. Each simulator is responsible for dispatching messages to the other simulators to convey the information they need to know about its actions. Conversely, each simulator is responsible for receiving, interpreting, and responding properly to messages received from other simulators.

A significant advantage of this distributed simulation approach is that as the simulation network is expanded, each new simulator brings with it all of the computational resources necessary to support itself. This means that adding new simulators does not (generally) involve hardware modifications to simulators already on the network.

Each simulator is, of course, responsible for maintaining a detailed model of its own state, including, for example, engine power, thrust, and fuel consumption; aerodynamic forces or terrain forces; weapon system computers, etc. Each simulator also maintains a simple "dead reckoning" model of the state of every other simulator on the network that is within possible interaction range. In essence, this involves extrapolating the last reported position of each other vehicle, based on its last reported velocity vector, until such time as a new state update message is received.

Each simulator is also responsible for sending out state update messages whenever it changes course or speed. To do this, each simulator must maintain, in addition to its "high-fidelity" model, a dead reckoning model that corresponds to the model that other simulators are maintaining of its state. In essence, after each update of its high-fidelity model, the simulator compares its exact state with that of the dead reckoning model and transmits a state update message only when a significant discrepancy has accumulated.

The state update message contains the essential "externally visible" information that the other simulators will need to paint an accurate picture of the sending vehicle on their screens: the vehicle's location and orientation, with six degrees of freedom (or more, if it has an independently adjustable turret, gun tube, or other feature visible at a distance), its velocity vector components, and whether it is currently producing smoke, a dust column, a muzzle blast, or other visual effects.

This approach leads to a variable update rate that will differ from one simulator to another at any given time. Each simulator transmits state update information only when necessary. The principal motivation is, of course, to minimize network message traffic and hence the amount of incoming information that other simulators must process.

SIMNET Protocols

Simulation protocols contain formal procedures for communicating data among the various components of the SIMNET network. One can think of them as a set of messages, with explicit rules about what data elements each message must contain, the conditions under which the message must be sent and to whom, the responses that other components are required to make upon receipt of the message, and so forth. A full discussion of the protocols is beyond the scope of this report, but the brief list of the most common messages in Table B.1 illustrates the protocols' scope.

These protocols are described in *The SIMNET Network and Protocols,* Bolt, Berenak, and Newman Corporation, BBN Report No. 7102, July 1989. This report is being used as the basis for proposed DoD-wide standards for interoperability of defense simulations.

Network Communications

Currently, Ethernet™ LAN technology is being used for all LAN connections, and standard 56-kilobit per second dial-up lines or T1 dedicated communications circuits are being used to link together the WAN. The SIMNET protocols do not depend on any particular LAN or WAN technology, as long as certain requirements for datagram delivery and transport delay are met. Both token-ring networks and fiber-optic networks are being explored for other applications.

Current Simulations

Currently, there are three categories of man-in-the-loop simulations in SIMNET. These categories differ in the level of detail with which the simulation is controlled.

Manned Simulators

In fully manned simulations, a full complement of crew members manipulate their own controls and observe outside phenomena from their out-the-window views or their onboard sensor views. For this category of simulators, the simulation code is designed to represent the characteristics of the vehicle and its various onboard systems, but makes no assumptions about the capabilities or decisions of the crew.

Table B.1

SIMNET Protocols

Message	From
Activate	Sent by a SIMNET Management, Command, and Control (MCC) system to a simulator, causing the simulator to become active.
Activating	Returned by a simulator as an acknowledgment of an Activate message.
Deactivate	Broadcast by an MCC system to discontinue the simulation of a vehicle.
Vehicle Appearance	Broadcast by a simulator, describing the location and appearance of its simulated vehicle. It is also broadcast by an MCC system to describe a vehicle it simulates.
Vehicle Status	Broadcast by a simulator to checkpoint the maintenance and supplies status of its simulated vehicle.
Fire	Broadcast by a simulator or an MCC system when ammunition is fired.
Impact	Broadcast by a simulator when its vehicle hits another vehicle (or the ground) with direct fire.
Indirect Fire	Broadcast by an MCC system when rounds impact from the howitzers and mortars it simulates, or when bombs detonate from the air strikes it simulates.
Collision	Broadcast by a simulator to indicate that its vehicle has collided with another.
Service Request	Broadcast by a simulator whenever its vehicle is capable of accepting supplies or repairs from a nearby combat service support truck.
Resupply Offer	Sent by an MCC system to a simulator, offering that simulator's vehicle some quantity of supplies.
Resupply Received	Sent by a simulator to an MCC system after having accepted supplies offered by a Resupply Offer message.
Repair	Sent by an MCC system to a simulator, describing a repair that has been completed on that simulator's vehicle.
Repaired	Sent by a simulator to an MCC system, to acknowledge the receipt of a Repair message.

Automated Simulations

Automated simulations, on the other hand, are used when there is no desire to emulate, train, or study the behavior of individual crew members. Automated simulations are used for artillery and mortar fire, and for some kinds of resupply functions. Here, the man-in-the-loop controller is acting as a resource allocator, receiving requests for fire support or logistics support over the radio, and deciding where to dispatch or deploy the limited set of resources he controls.

No attempt is made to model the control behavior of the fuel truck driver or the calculations carried out by the gun crew. We are concerned only with realistically representing the *effects* of these activities on the battlefield, not their details.

Semi-Automated Simulations

Semi-automated forces (SAFOR) are an intermediate case. Here, we do want to represent in some detail the behavior of individual vehicle crews and small units, but with a small number of individuals controlling multiple vehicles. To achieve this, it is necessary to model in considerable detail the control and tactical decisions of the crew as well as the dynamics of the vehicle.

The SIMNET SAFOR are currently generated by a three-component system. First, there is the SAFOR simulation host, which is responsible for receiving and sending messages on the SIMNET LAN in exactly the same format as that used by manned vehicle simulators. In addition to maintaining the dynamic state of each vehicle for which it is responsible, the simulation host performs local terrain navigation, route following, and formation keeping. It also continuously recomputes intervisibility among nearby vehicles, and executes target acquisition and firing procedures in accordance with the SAFOR commander's guidance. It is also responsible for composing and dispatching contact reports and situation reports under various circumstances, so that the SAFOR commander and his simulated staff can tell what's going on.

The second SAFOR element is the commander's workstation, which contains a map display, a communications log, and a operations order section by which the human commander instructs his units. This workstation is responsible for fulfilling many of the functions of the commander's staff and subordinate unit commanders. That is, it will make "reasonable" inferences about the commander's intent so that he does not need to issue orders at an unrealistic level of detail. This requirement implies that the workstation have at least a limited capability to predict the consequences of alternative actions in order to decide what the commander probably meant.

There are, obviously, limits to the extent to which we can make these inferences, so there are several categories of decisions that are reserved for the third element of the SAFOR system, the human commander. Only the commander can change a mission (from attack to defend, for example), or order the firing or lifting of final protective fires. Moreover, the simulated subordinate unit commanders are required to inform their human commander whenever they are about to become "decisively engaged," that is, to take an action that may prove irreversible.

SAFOR have been generated to represent both Red and Blue versions of each of the manned vehicle simulations we have developed: tanks, infantry fighting vehicles, helicopters, close-air support aircraft, and air defense weapons. More SAFOR vehicles will be added as the battlefield expands.

Data Collection and Analysis

A useful by-product of the distributed simulation protocols is that the various messages can be time-stamped and recorded by a host computer on the simulation network, and then played back onto the network at a later time. An observer can use any simulator to "time travel" throughout the battlefield, seeing exactly what he would have seen if he had been in that position when the battle first took place. His simulator will not be able to tell whether the messages from other vehicles were prerecorded or are being generated in real time. It should be noted, however, that the observer cannot change the course of the battle he is observing—all of the actions of the vehicles are predetermined by what they did when the battle was first recorded.

The data collected by the SIMNET data logger can then be used for off-line statistical analyses, using several software packages developed or adapted for the purpose.

Stealth Vehicle

A special category of simulator is the "stealth" vehicle, which can be initialized so that it does not broadcast appearance messages to the other simulators. This simulator can be used by the unit commander or an experimental analyst to observe the conduct of an exercise without affecting it. The stealth vehicle is equipped with either generic ground vehicle dynamics, so that it follows the terrain, or "flying carpet" dynamics, so that it can move freely about the battlefield in any direction and will remain hovering whenever the controls are released.

The stealth vehicle can also be attached to any selected vehicle, so that it can follow it about the battlefield without operator intervention.

Electro-Optical and Thermal Sensors

For some of the SIMNET applications, we developed a fairly simple simulation of electro-optical and thermal sensors. In effect, we employed alternative sets of vehicle models to represent approximately how they would appear on the sensor displays, and used alternative terrain colors and textures to represent the appearance of the surrounding terrain. While these simulations were adequate for our initial purposes, they are not able to support the kinds of simulations envisioned in the CSRDF/SIMNET study.

Radar Simulation

A fairly detailed digital radar simulation has been developed for use by the air defense simulators. This radar simulator is capable of being used while in motion, and continually recalculates the line-of-sight to each nearby target to see whether it is masked by terrain or other objects. Knowing the type of target, the range, the azimuth, and the relative orientation of each potentially observable object, the radar simulator can use downloadable look-up tables to determine the probability that each object would be tracked and displayed on its screen.

This radar simulator also broadcasts appearance messages of its own, indicating what sector it is illuminating at what time. These appearance messages can be used by other simulators to detect the simulated radiation, triggering a radar warning receiver or perhaps guiding an energy-seeking missile toward the radar's antenna.

Missile Models

Several missile models have been developed for use in various simulations. Briefly, they include those that are guided by a remote device (such as a tube-launched, optically tracked, wire-guided [TOW] missile being guided by a human operator), those that use onboard sensors and guidance systems to home in on energy of some kind being radiated by a target, and those that use a third-party designator to "paint" the target. Each of these models can be modified to represent the characteristics of a variety of guided missiles.

Radio Communications

A significant capability developed under SIMNET is the radio communication propagation model, which can support a wide range of command and control functions. Radio simulators broadcast communications data on the network in much the same way as the vehicle simulators broadcast vehicle appearance data. That is, whenever a transmitter is radiating, the radio simulator performs an analog-to-digital transformation on the voice signal, digitally compresses it to conserve bandwidth, and adds a message header that indicates the location of the transmitter, its frequency, power, antenna direction, and so forth.

Radio receiver simulators receive these packets and quickly filter them to determine which are on the frequency to which they are tuned and are within potential receiving range. They then carry out fairly detailed signal propagation calculations, including both range effects and first-order diffractions over large terrain objects, as well as applying models of their own signal capture characteristics. Only the "winning" signal (if any) will be converted back into an analog signal and played into the headsets of the crew. This signal could of course be a jammer rather than the desired signal. The jamming could be deliberate or inadvertent.

C. Crew Station Research and Development Facility[1]

The Crew Station Research Development Facility (CSRDF) resulted from a high-priority joint Army-NASA development project that was initiated in 1985. The CSRDF serves as a full combat mission simulator with emphasis on research into mission equipment and crew vehicle interfaces for Light Helicopter (LH) type weapon systems. Primary design requirements for the CSRDF are

- support the Army Composite Mission Scenario for research, development, and evaluation;
- provide a high-fidelity rotorcraft simulation;
- provide reconfigurability for the main simulation vehicle and auxiliary/support vehicles; and
- support a wide range of threat types and capabilities.

The basic combat environment designed for the CSRDF has been approved by the Army Aviation Center and has been configured to support a wide range of environments. The crew stations of the CSRDF are glass cockpits that faithfully represent all crew station interfaces for night adverse weather missions. All subsystems have the appropriate levels of interaction with other subsystems and combatants; fields of view, intervisibility, and countermeasures at all wavelengths are modeled.

The CSRDF provides limited capabilities for man-in-the-loop war gaming and evaluation. A total of 100 ground elements and 11 air elements are simulated. The air elements can consist of a Scout/Attack team, utility rotorcraft, or threat rotorcraft. All ground elements are either automated or scripted, while air elements are controlled through the use of three interactive graphics workstations that provide the capability to control up to four aircraft each.

The CSRDF is a traditional simulator consisting of a central computer complex that drives (1) a complex research cockpit with visual displays and (2) auxiliary

[1]This appendix was prepared by the U.S. Army Aviation Research and Technology Agency (ARTA) and the National Aeronautics and Space Administration (NASA) at the Ames Research Laboratory. Authors were Terrence Gossett, Edward Huff, and James Voorhees.

workstations that provide simulated interaction with the research cockpit. See Figure C.1.

The crew station cockpit contains two stations, with the front seat normally used by the pilot and the back seat by the battle captain. All functions, including flight controls, can be assigned to either seat so that one individual can act as both pilot and battle captain to evaluate workload factors. A full suite of controls and displays, simulated weapons, sensors, and survivability equipment is provided in the CSRDF as is a voice recognizer. The latter may be used to test the functionality of using voice commands for controlling simulated sensors and other equipment. All displays in the cockpit are fully reconfigurable, and the pilot is provided with a Fiber Optic Helmet Mounted Display (FOHMD) that provides a 67° × 127° (V × H) field-of-view and an unlimited field-of-regard.

The CSRDF also has three team stations implemented in general-purpose graphics workstations that are capable of controlling from one to four helicopters each, and a "White station" that may serve as a C^{3}I station or a command post. A voice disguiser is included at the White station so that one individual can operate as a variety of communicators.

RANDMR340-C.1

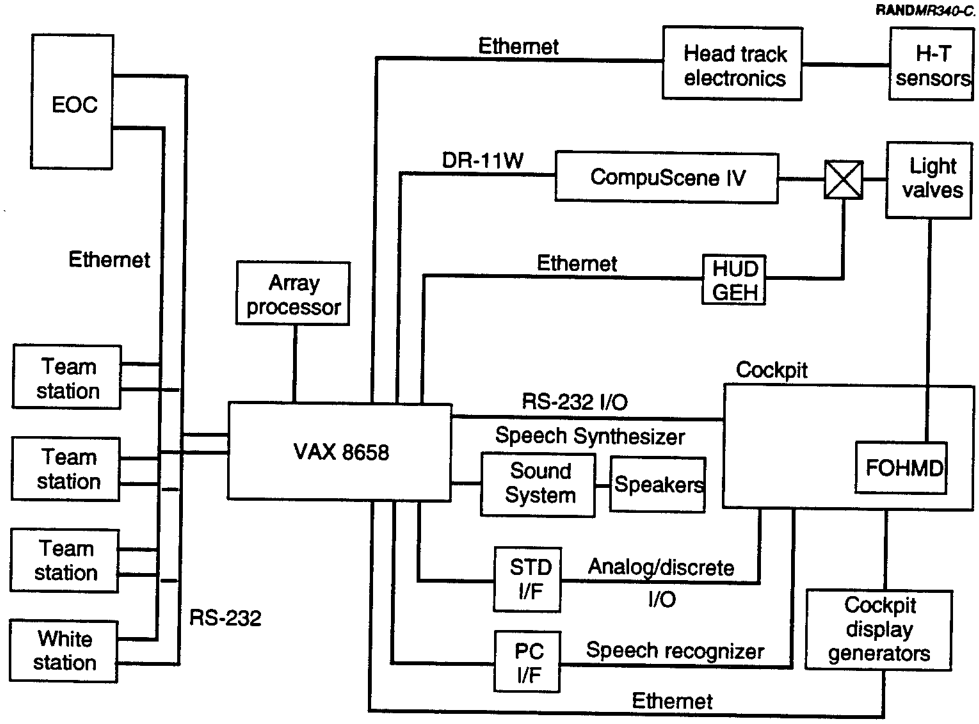

Figure C.1—CSRDF Architecture

The CSRDF incorporates an experimenter/operator console and a data analysis and reduction facility that allow development of scenarios, control of the simulation, and analysis of the simulation results. Various interfaces are provided for the researcher so that he or she may analyze the subject's action and interact with the scenario in various ways (such as the insertion of fault modes in the primary research platform). All kinematics (except the team station kinematics), weapons interaction, sensor interactions, countermeasure interactions, and other functions are calculated primarily in the host computer complex. The team stations are connected to the host computer via standard, individual Ethernet connections.

The battlefield environment of the CSRDF consists of Defense Mapping Agency (DMA) Digital Terrain Elevation Data and Digital Feature Analysis Data (DTED and DFAD) that have been modified and enhanced to provide compatibility with CSRDF graphics generation equipment. The terrain and feature data have also been modified to provide map displays at the crew station, team stations, and the researcher's console.

Out-the-window (OTW) views and sensor displays can be presented in a wide range of environmental conditions that include day, dusk, and night with a variety of weather effects. Eye points within the computer-generated image systems can be located to represent the view from any particular sensor, such as from the mast or the nose of the aircraft. A sophisticated sound simulation system is included in the CSRDF that provides realism for the research subject. Sounds simulated include own-ship rotor and engine noise and weapons effects.

Graphic displays available in the crew station include color, day/dusk/night OTW, white or black hot FLIR, and day television (DTV).

All flight management and mission management crew-vehicle interfaces specified in the LH demonstration/validation contract are represented in the CSRDF.

In addition to the primary research platform and team stations listed previously, a maximum of 100 computer-controlled interactive threats may be included in the scenario at any time. These interactive threats can be automated or scripted and include a wide variety of friendly aircraft, threat aircraft, tanks, anti-aircraft artillery (AAAs), surface-to-air missiles (SAMs), and various other wheeled, tracked, or fixed-position entities. All players in the gaming area are provided with a full set of weapons, sensors, and survivability equipment appropriate to their weapon system.

Players may be reconfigured, replenished, or otherwise modified during the simulated mission to present the battle commander with realistic decisions, workload, and battle actions.

The emplacement, command and control, tactics, and scenario for each of the combatants in the simulated mission are carefully structured to reflect Blue/Red doctrine.

As stated earlier, the CSRDF is a traditional simulator that uses a centralized computer complex to calculate vehicle kinematics, sensor interactions, and all other data necessary to support the simulation. The central computer complex (CCC) includes a high-speed array processor to process an own-ship 10-element rotor blade model and certain stabilization systems. The CCC also hosts all own-ship fuselage, auxiliaries, and mission equipment package (MEP) electronics models as well as weapon, sensor, and countermeasure models for all gaming area players.

The CCC is connected to various display processors (graphics workstations) in the team stations and crew station by a number of Ethernet LANs. Sufficient data for each workstation to generate a map or a view are passed from the CCC common database during synchronous data transmission periods. Thirteen workstations are required to process all displays present in the CSRDF.

A high-speed parallel interface is provided to transfer data between the CCC and the crew station visual image generator (IG), a General Electric CompuScene IV. Pilot head orientation with respect to the gaming area is calculated (and predicted through the use of accelerometers) and passed to the IG to indicate scene orientation. The IG then assembles the view polygon list, modifies it with respect to environment and sensor in use, and produces the scene video. The video consists of two wide fields-of-view and two high-resolution views that are optically combined. Two separate outputs are piped over fiber optic bundles to the FOHMD. Head-up display video is provided through the optical path.

In summary, the CSRDF provides a wide variety of tools that may be used in the acquisition and decisionmaking processes. The high-fidelity simulation models and realistic crew station environment used in the CSRDF ensure that data collected and analyzed will be valid for a broad range of uses. The CSRDF does, however, have a number of limitations:

1. A limited number of combatants and types of combatants. The initial design of the CSRDF limits the overall scenario to 111 gaming area entities.

2. Sparseness of the terrain database. Notwithstanding the state-of-the-art in computer image generation, the number of gaming area cultural features (trees, buildings, etc.) is not large.
3. A limited number of human operator-controlled combatants. Fog-of-battle issues may not be fully or properly addressed, hence missing the richness, robustness, and capriciousness of human operations.
4. Lack of special sensors. During design of the CSRDF it was decided to not simulate specific equipment (such as specialized ELINT or COMINT).
5. Single pilot day/two pilots at night. Cost constraints limited the number of daylight OTW displays to one crew station, although both crew station positions can display FLIR images.
6. Unclassified facility. During initial design, it was decided that including classified parameters in the characteristic databases did not warrant the additional cost. Thus, models use approximations of vehicle, countermeasures, sensors, and weapon characteristics that have been derived from actual Blue/Red systems. Although models and parametric data have been fully approved by cognizant simulation activities, some finer points of battlefield interactions may not be addressed.

References

Kerr, Andrew, *Aeromechanics and Man-Machine Integration Technology Opportunities for Rotorcraft of the 1990s and Beyond*, AIAA-89-2065, August 1989.

Hubbard, Robert, and Darryl Caldwell, *LHX Helicopter DEM/VAL Program*, AIAA 89-2115, August 1989.

Thorpe, Jack A., USAF/DARPA, "Warfighting with SIMNET—A Report from the Front," IITSC, December 1988.

Bibliography

Anderson, N. H., "Functional Measurement and Psychophysical Judgment," *Psychological Review*, 77, 1970, pp. 153–170.

Anderson, N. H., *Foundations of Information Integration Theory*, Academic Press, New York, 1981.

Birnbaum, M. H., "The Nonadditivity of Personality Impressions," *Journal of Experimental Psychology*, 102, 1974, pp. 543–561 (monograph).

Birnbaum, M. H., "On Rescaling Data to Fit the Model and Concluding that the Model Fits: A Note on Monotonic Transformation," *Perception and Psychophysics*, 1982, pp. 293–296.

Birnbaum, M. H., "Scale convergence and psychophysical laws," in H. G. Geissler, M. H. Mueller, and W. Prinz (eds.), *Psychophysical Explorations of Mental Structures*, Hogrefe & Huber, Toronto, 1990.

Birnbaum, M. H., and C. T. Veit, "Scale Convergence as a Criterion for Rescaling: Information Integration with Difference, Ratio, and Averaging Tasks," *Perception and Psychophysics*, 1974b, 15, pp. 7–15.

Birnbaum, M. H., and C. T. Veit, "Scale-Free Tests of an Averaging Model for the Size-Weight Illusion," *Perception and Psychophysics*, 1974a, 16, pp. 276–282.

Green, D. M., and J. A. Swets, *Signal Detection Theory and Psychophysics*, John Wiley and Sons, Inc., New York, 1966.

Hubbard, Robert, and Darryl Caldwell, *LHX Helicopter DEM/VAL Program*, AIAA 89-2115, August 1989.

Kerr, Andrew, *Aeromechanics and Man-Machine Integration Technology Opportunities for Rotorcraft of the 1990's and Beyond*, AIAA-89-2065, August 1989.

Krantz, D. H., and A. Tversky, "Conjoint-Measurement Analysis of Composition Rules in Psychology," *Psychological Review*, 1971, 78, pp. 151–169.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky, *Foundations of Measurement*, Academic Press, New York, 1971.

Thorpe, Jack A., USAF/DARPA, "Warfighting with SIMNET—A Report from the Front," IITSC, December 1988.

Veit, C. T., "Ratio and Subtractive Processes in Psychophysical Judgment," *Journal of Experimental Psychology*, Vol. 107, 1978, 1, pp. 81–107.

Veit, C. T., and M. Callero, *The Subjective Transfer Function Approach to Complex System Analysis*, RAND, R-2719-AF, 1981.

Veit, C. T., *Validating Judgments and Decisions for Combat Simulations: Issues and Procedures*, RAND (forthcoming).

Veit, C. T., M. Callero, and B. J. Rose, *Introduction to the Subjective Transfer Function Approach to Analyzing Systems*, RAND, R-3021-AF, 1984.