VERIFICATION & VALIDATION IN MILITARY SIMULATIONS

Dean S. Hartley III

Data Systems Research and Development
1099 Commerce Park
Oak Ridge, Tennessee 37830, U.S.A.

ABSTRACT

The verification and validation (V&V) of military simulations has many similarities to the V&V of simulations as often discussed in academe; however, significant differences exist. Military simulations are often large, extremely complex production systems, in which the initial conditions provide the driving force and the transients are the elements of interest. The difficulties of performing the V&V and the lack of understanding of the realities being modeled have led to fears that the costs of V&V would greatly outweigh any benefits. Despite these fears, the current Defense Department requirement for the verification, validation and accreditation (VV&A) of simulations and the desires of modelers and users for useful simulations are generating a considerable body of experience in the actual V&V of military simulations.

1 BACKGROUND

Discrete event and time stepped simulations have been the foundations for the major computer tools for analyzing combat since the 1960s (ATLAS, IDAGAM, VECTOR, CEM, TACWAR, THUNDER, VIC, CASTFOREM, Janus). [See Appendix for model data.] Simulations have also been the foundations for the major computerized combat training tools (Janus, CBS, RESA, MTWAS, AWSIM, JTLS). Major efforts are currently under way to produce a new generation of tools, JWARS for analysis and JSIMS for training. Decisions (e.g., weapons procurement) and actions (e.g., training doctrine) involving billions of dollars have been and will continue to be influenced by these simulations. Realizing this, the Department of Defense (DoD) has made the validation, verification and accreditation (VV&A) of these simulations an official requirement (DODD 5000.1, DODI 5000.61).

DoD purchasers of combat simulations (and the simulation builders) have always wanted these tools to be as accurate as possible; however, the allocation of resources often left little time and money for V&V. In addition, the V&V of combat models presents technical problems that differ from those often discussed in V&V of general simulations. The magnitude of these technical problems has led some to conclude that V&V must be prohibitively expensive, if done thoroughly.

This paper will address V&V techniques and problems, applications and cost, and the philosophical issue of how much is enough.

2 TECHNIQUES: DEFINITIONS, METHODOLOGY AND PROCESS

Informally, VV&A may be defined as:

Verification: does it do what they said? (Build the model right.)
Validation: is it right? (Build the right model.)
Accreditation: is it good enough for government work?

Williams and Sikora (1992) give the definitions developed by the Military Operations Research Society (MORS) and adopted by the Department of Defense (DODI 5000.61, differences shown in brackets). These definitions are related in 1, including the reference to the fuzzy, proxy for the real world that must suffice for military simulations.

VERIFICATION: The process of determining that a model implementation accurately represents the developer's conceptual description and specifications.

VALIDATION: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

ACCREDITATION: An official determination [certification] that a model [, simulation, or federation] is acceptable [for use] for a specific purpose.
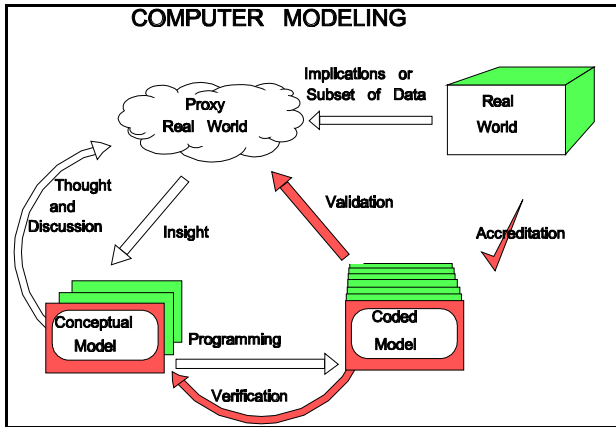
Figure 1: V&V Relationships

Averill Law and David Kelton (1982) wrote, "One of the most important problems facing a real-world simulator is that of trying to determine whether a simulation model is an accurate representation of the actual system being studied, but a review of the validation literature indicates that relatively little has been written on this subject. Furthermore, what has been written is often philosophical in nature rather than in the form of practical recommendations."

Several authors have worked to improve this situation. Methodologies have been defined for V&V of general simulations: Naylor and Finger (1967), Schrank and Holt (1967), Kostelski et al. (1987), Balci

(1989), Gross (1996), and especially Sargent (most recently 1996). Additionally, Whitner and Balci (1989) and Balci (1996) have developed taxonomies of the V&V methodologies.   Some authors have discussed the procedural aspects, including the need to begin V&V at the beginning of the simulation project and the need for V&V of the data (Carson 1989).

Francis Hoeber (1981) wrote, "As in most cases large [military] models, validation of the [particular model] turns out to be virtually impossible at a level that will satisfy critics as well as users." "...the Air Force appears to believe that the [model] has conceptual, or face, validity." Sikora and Williams (1997) find the situation to be considerably better now.

There has been work done on the methodologies appropriate for V&V of military simulations: Banks (1989), Gaver (1992a, 1992b), Giadrosich (1992), Henderson (1992), Metzger (1992), and Chew (1997). Cynamon (1992) has addressed the parts that documentation and configuration management play in V&V. Seglie and Sanders (1992) discuss accreditation. Paul Davis (1992) has done an admirable job of pulling much of this together and producing a taxonomy of military VV&A methodologies.   Hartley and Whitley (1996) and Stanley (1997) add the dimension of V&V of data to the picture. 2 builds on Davis' taxonomy, adding the data dimension and making other minor modifications.
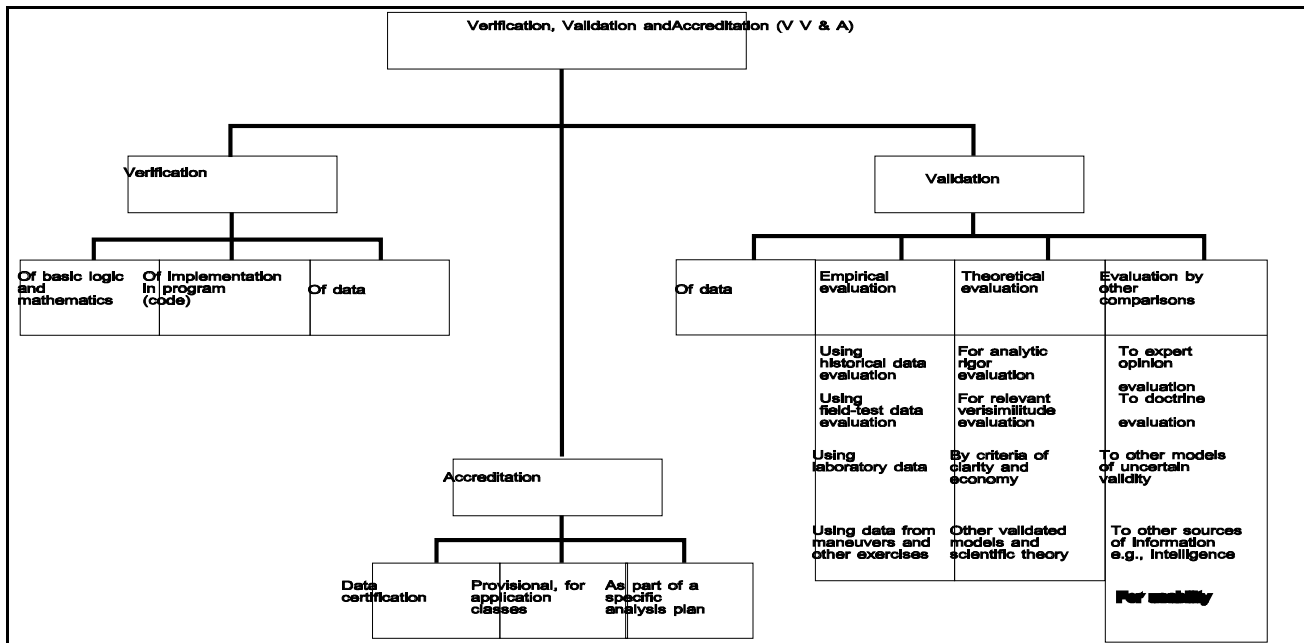

Figure 2: VV&A Taxonomy

It would seem obvious that V&V started at the beginning of a simulation project would find errors sooner, reducing the cost of corrections (3). However, V&V is often not considered until a model is complete, or nearly so.
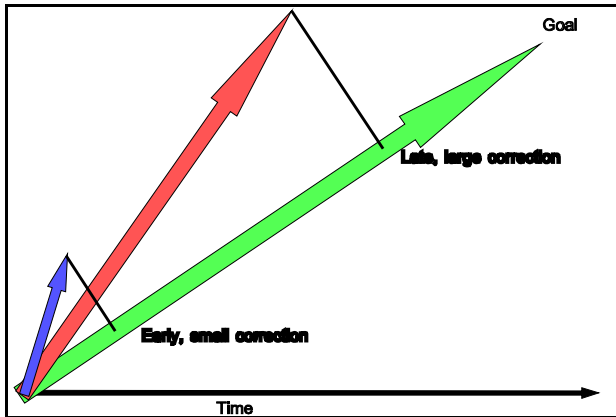

Figure 3: Early Corrections vs Late Corrections

Except for V&V performed on legacy models, the current DoD regulations yield regular corrections over time, referred to as continuous V&V. Properly applied, these corrections should result in a simulation that achieves its goal, as shown in 4. This procedure is also sometimes referred to as a "build-test-build" procedure.
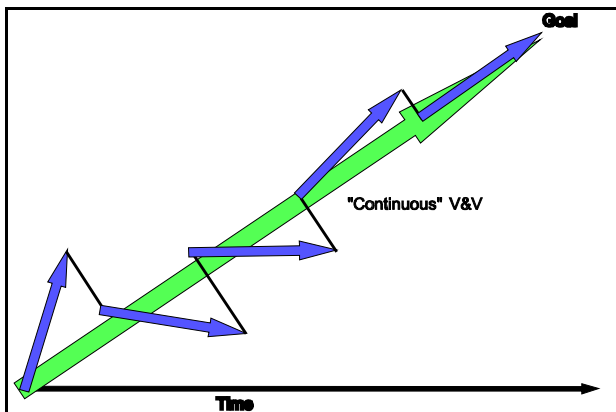

Figure 4: Continuous V&V

Independent V&V (IV&V) is added to the process through a series of tests of the internal V&V procedures and (possible) additional corrections, as shown in 5. The timing of the IV&V activities may coincide with major project milestones.
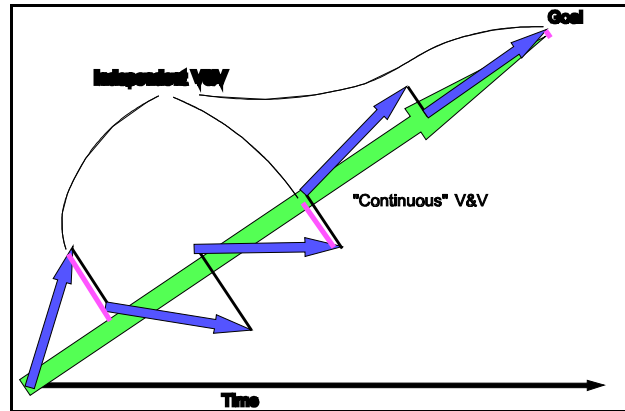

Figure 5: Adding IV&V

3 PROBLEMS

One area of concern to general simulation V&V is the concept of achieving steady-state in a simulation (Gross 1996). Unfortunately, combat simulations are interested in the transient effects produced by initial conditions, not some steady state condition. In addition, the sheer size of military simulations proliferates logic paths exponentially, making exhaustive path testing impossible. Further, some military simulations employ human-in-the-loop, or human decision-making, as part of the simulation system. This makes repeatability impossible.

The recent introduction of confederations of simulations introduced (or re-introduced) the problem of compatibility of differing modeling assumptions and techniques. Confederation problems include:

The rapid pace of innovation, multiple configuration control schedules, and cost of joint tests reduces testing.

A flexible confederation of connected entities (plug & play) obscures the identity of the model: A+B+C, A+C, A+B+D, etc.

When an error is found, whose problem is it?

Who pays for V&V when the separate simulations are owned by different organizations?

A simplified discussion of a real case helps illustrate the problem of confederated simulations. Suppose a confederation includes three simulations, Air, Ground, & Sea, each having been VV&Aed and found to be perfect. The Air and Sea simulations use a "hit is a kill" model, while the Ground simulation uses separate probability of hit (pH) and probability of kill-given-a-hit (pK|H) values. In this confederation, the adjudication is divided so that the simulation containing the attacking weapon defines the trajectory of weapon (whether there is a hit) and the simulation containing the attacked platform determines the

damage. Consider the case where the attacking weapon (W_A) is in the Air simulation, helicopter (H_G) is in the Ground simulation, and an identical helicopter (H_S) is in the Sea simulation. The Air simulation's W_A determines an X% probability of hit, therefore X% of P_S are killed, using the Sea simulation logic, but X% times pK|H (less than X%) of P_G are killed, using the Ground simulation logic. How do you set X (remembering that the Ground simulation must be able to use its own logic to fire on helicopters in the Ground simulation, also)? Despite that assumption that all three simulations are valid, the confederation produces a model of reality that is invalid. 6 illustrates the concept of joining parts that fit mechanically, but have no overall validity.
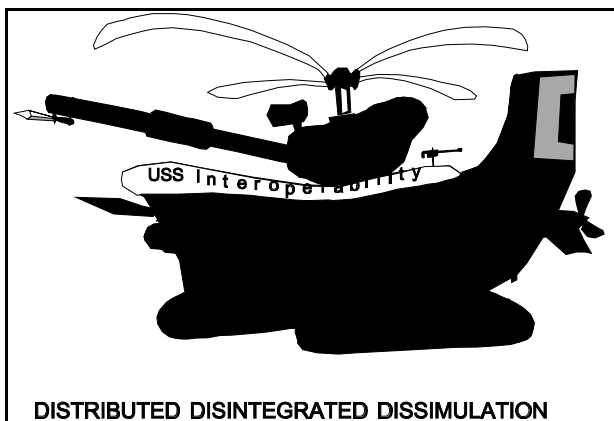

Figure 6: Accredited Parts, Invalid Whole

　　Perhaps the most critical problem for military simulations is the fact that the structures of combat processes and the interactions of these processes are essentially unknown. For example, most combat simulations use the Lanchester square & linear models for attrition processes, despite evidence that these models don't match historical combat (Hartley and Helmbold 1995 and Hartley 1995b). Most combat simulations couple force movement to the attrition equations, despite the fact that no evidence supports such a coupling (Helmbold 1995). Despite these problems, there is still a correspondence between historical combat and the results of combat models (Whitley 1991) - but we don't know why.

5 APPLICATIONS OF V&V

Reports of the actual performance of V&V on military combat simulations are sparse, but appear to be increasing in number and in generality of model coverage. Hartley (1975) reported that the TAC CONTENDER model was not a global optimizer, as claimed. Tufarolo reported on

the components of V&V that were included in the development of the JTLS simulation. Hartley et al. (1989) reported on efforts to perform sensitivity analysis on JTLS. Hartley, Quillinan and Kruse (1990) reported on V&V of certain algorithms of the SIMNET system of simulations. Hartley, Radford and Snyder (1991) reported on V&V of SIMNET with respect to its domain of usefulness.

　　More recent work involves V&V of the complete simulation. Hartley et al. (1994) performed V&V on a model at the very beginning of its life cycle. Muessig (1997) reports on the complete V&V of a number of legacy weapons engineering-level simulations. Youngblood (1997) reports on the integrated V&V process being used in the creation of JWARS.

6 COSTS

Model development (as it actually occurs) includes both model creation (concept and coding) and V&V. The upper left quadrant of 7 shows an estimated split wherein creation costs exceed V&V costs; however, the time/cost contributions are comparable. The upper right quadrant shows an estimated split of these V&V costs between verification and validation. Verification necessarily predominates in the V&V of model development. These estimates are based on personal experience and reports such as Tufarolo's (1986). The bottom half of 7 estimates the impact of adding IV&V to model development. The increase shown is on the order of 15%. Lewis (1997) describes VV&A costs as ranging from 5% to 17.5% (his "V&V" appears to correspond to IV&V as defined here). Muessig (1997) describes V&V of six legacy models (which corresponds to IV&V here) and gives the total IV&V cost at just under 10% of the total model development costs. The split in IV&V costs, in the lower right quadrant, is based on including spot checks of verification and giving heavy emphasis to validation, thus producing a better total balance between verification and validation.
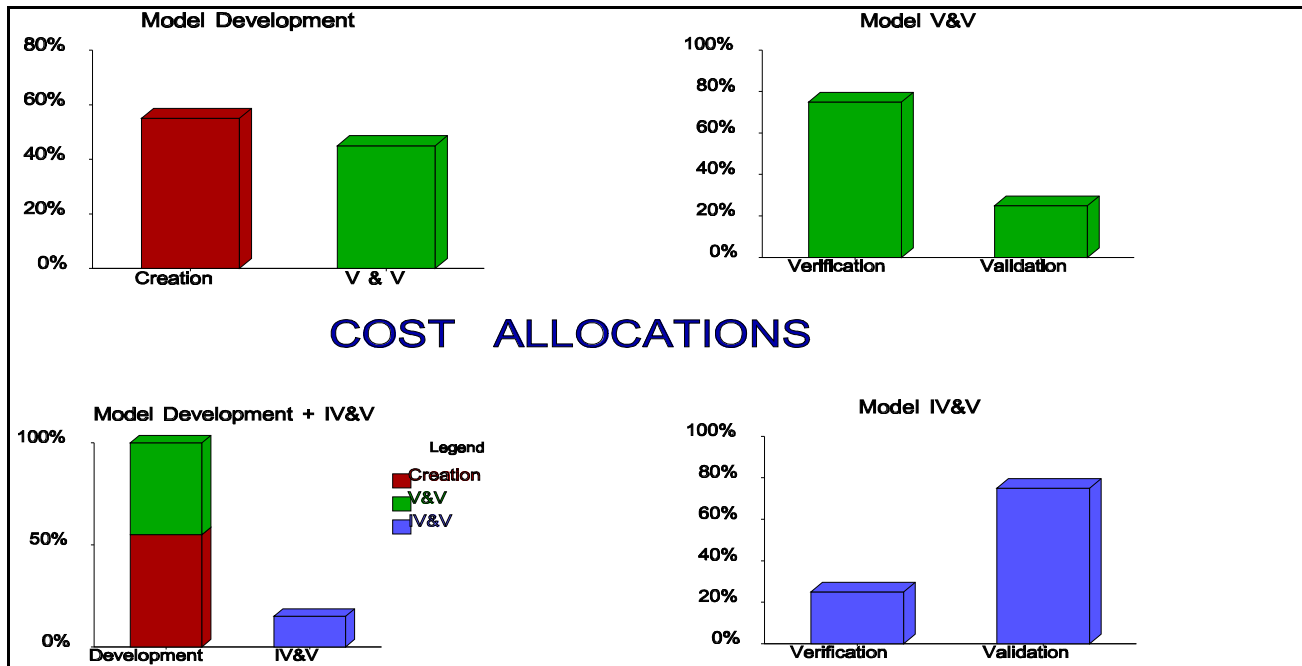
COST ALLOCATIONS

Figure 7: Military Simulation Cost Components

7 PHILOSOPHY

Francis Kapper (1981) wrote, "The most appropriate and valid objectives for using war games and simulations within the DoD context are to: better understand complex phenomena, identify problems, evaluate alternatives, gain new insights, and broaden one's perspectives. The least valid or appropriate objectives for using war games and simulations are to predict combat/crisis outcomes or control broad and highly complex programs." His statement was based on an understanding of the level of validity achievable by military simulations and the implications on the proper uses of those simulations.

The intended use of model is the proper driver for VV&A. The concept of differing uses means differing emphasis of functional models and differing precision requirements. In defining V&V, remember the Accreditation process (that focuses on use) and the cost/benefit of the V&V.

If the use is training, the goal is enhanced training value with no negative training. In training, some process detail may be less important than some visual presentation (as compared to use in development or analysis). On the other hand, it is possible that increasing the fidelity of the visual representation in virtual reality simulations may increase the subliminal acceptance of incorrect assumptions.

Things may be worse than they seem...
If your business is combat...
and there isn't any around...
and you're bright and ambitious...
...you learn what and where you can!

What are we teaching our future generals that isn't so? (Hartley 1995a)

In development, the use of the complete spectrum of possible results may be more important than in training, where exposure to the more unlikely possibilities may be undesirable.

In analysis, the simulation must differentiate among the alternatives being analyzed, but need not differentiate elsewhere.

In deciding how much is enough, a V&V plan is required. It should be based on the intended use. For example, in training, the following provides a template:

Pick the operational training objectives that apply.
Test the model against those.
For each training objective:
    Does the model address the objective?
    How well?
    What are potential negative training points?

Once the plan is defined and has been determined to be comprehensive, avoid additional issues.

8 SUMMARY

We can never be confident of combat simulation validity until we are confident that we understand combat. Research is required and, until we understand combat, we must regard any validation as contingent. 8 represents the panels in the Sunday comic strips in which the reader is supposed to find six errors. However, in V&V of military simulations, the problem is harder. We don't know how many errors there are, or if there are any.
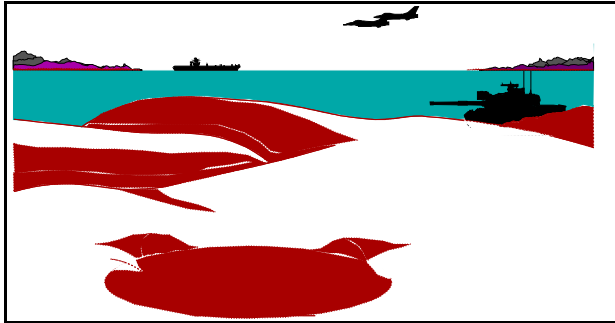

Figure 8: What is wrong with this picture?

Despite the problems, the importance of the V&V of military simulations remains. Simulations influence billions of dollars of investments and create the possibility of negative training for our forces. This importance is magnified today because our forces are small and in fast wars we can't afford mistakes, because we don't have the time and resources to recover. Well done V&V of our military simulations will reduce the number and magnitude of our mistakes.

ACKNOWLEDGMENT

APPENDIX: MODELS

This sections gives a brief description of each of the campaign simulations mentioned in the text.

ATLAS - A Tactical, Logistical, and Air Simulation. Date Implemented: ~1969. Model Type: Analysis. Proponent: Research Analysis Corporation.

AWSIM - Air Warfare Simulation. Date Implemented: 1988. Model Type: Training and education. Proponent: HQ USAFE Warrior Preparation Center (WPC), Einsiedlerhof Air Station, Einsiedlerhof, Germany APO AE 09094-5000.

CASTFOREM - Combat Arms Task Force Engagement Model. Date Implemented: 1983. Model Type: Analysis. Proponent: TRADOC Analysis Command, White Sands Missile Range (TRAC-WSMR), White Sands Missile Range, NM 88002-5502.

CBS - Corps Battle Simulation, Version 1.3.5; formerly known as Joint Exercise Support System (JESS). Date Implemented: 1985. Model Type: Training wargame.

CEM - Concepts Evaluation Model. Date Implemented: 1974. Model Type: Analytical. Proponent: U.S. Army Concepts Analysis Agency.

IDAGAM - Institute for Defense Analyses (IDA) Ground Air Model. Date Implemented: ~1974. Model Type: Analytical. Proponent: IDA.

Janus. Date Implemented: 1978 (latest version June 1991). Model Type: Depending on the needs of its users, Janus is being used for both analysis and training and education. Proponent: Conflict Simulation Laboratory, Lawrence Livermore National Laboratory, P.O. Box 808 L-315, Livermore, CA 94550.

JSIMS - Joint Simulation System. Date Implemented: Under development. Model Type: Training. Proponent: Joint Simulation System Program Office.

JTLS - Joint Theater Level Simulation. Date Implemented: 1983, with continuous functional upgrade since then. Model Type: Analysis and training. Proponent: Joint Warfare Center, Hurlburt Field, FL 32544.

JWARS - Joint Warfare System. Date Implemented: Under development. Model Type: Analysis. Proponent: Office of the Director for Program Analysis and Evaluation (ODPA&E) and the Joint Staff (J-8).

RESA - Research, Evaluation, and Systems Analysis Facility, formerly Interim Battle Group Tactical Trainer (IBGTT). Date Implemented: 1982. Model Type: Analysis and training. Proponent: Naval Ocean Systems Center, San Diego, CA 92152.

TAC CONTENDER. Date Implemented: ~1970. Model Type: Analysis. Proponent: Air Force.

TACWAR - Tactical Warfare. Date Implemented: ~1978. Model Type: Analysis. Proponent: Joint Staff, Force Structure, Resource, and Assessment Directorate (J-8), Automation Support Division, Washington, DC 20318-8000.

THUNDER. Date Implemented: 1984. Model Type: Analysis. Proponent: Analytic Model: Air Force Studies and Analyses Agency (AFSAA/SAG), Pentagon, Rm 1D380, Washington, DC 20330-5420.

VECTOR. Date Implemented: ~1974 (two major revisions, latest in 1982). Model Type: Analysis. Proponent: Vector Research, Incorporated, PO Box 1506, Ann Arbor, MI 48106.

VIC - Vector In Commander. Date Implemented: 1986. Model Type: Analysis. Proponent: TRAC-WSMR, White Sands Missile Range, NM 88002-5502.

REFERENCES

Balci, Osman. 1986. "Credibility Assessment of Simulation Results." *Proceedings of the 1986 Winter Simulation Conference*, ed. J. Wilson, J. Henriksen and S. Roberts, 38-44.

Balci, Osman. 1996. "Verification, Validation and Testing of Models." *Encyclopedia of Operations Research and Management Science*, ed. Saul I. Gass & Carl M. Harris, 719-723. Boston, MA: Kluwer Academic Publishers.

Banks, Jerry. 1989. "Testing, Understanding, and Validating Complex Simulation Models." *Proceedings of the 1989 Winter Simulation Conference*, ed. E. MacNair, K. Musselman and P. Heidelberger, 549-551. Washington, DC.

Carson, John S., II. 1989. "Verification and Validation: A Consultant's Perspective." *Proceedings of the 1989 Winter Simulation Conference*, ed. E. MacNair, K. Musselman and P. Heidelberger, 552-558. Washington, DC.

Chew, Jennifer P. M. 1997. "Testing and Validation of Models & Simulations, *Phalanx*, Vol. 30, No. 2.

Cynamon, Joseph J. 1992. "The Basics." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, II 1-16. Alexandria, VA: MORS.

Davis, Paul K. 1992. "A Framework for Verification, Validation, and Accreditation." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, VI 1-39. Alexandria, VA: MORS.

DODD 5000.1 - Department of Defense Directive 5000.1.

DODI 5000.61 - Department of Defense Instruction 5000.61.

Gaver, Don P. 1992. "Face Validation and Face Validity." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, IV 21-22. Alexandria, VA: MORS.

Gaver, Don P. 1992. "Sensitivity Study of a Simulation Model." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, IV 23. Alexandria, VA: MORS.

Giadrosich, Donald. 1992. "Validating Models and Simulations." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, IV 1-16. Alexandria, VA: MORS.

Gross, Donald. 1996. "Simulation of Discrete-Event Stochastic Systems." *Encyclopedia of Operations Research and Management Science*, ed. Saul I. Gass & Carl M. Harris, 626-633. Boston, MA: Kluwer Academic Publishers.

Hartley, Dean S., III. 1975. *An Examination of a Distribution of Tac Contender Solutions*, TM 101-75, AD A014 426. Washington, DC: NMCSSC.

Hartley, D. S., III. 1995. "A Mathematical Model of Attrition Data." *Warfare Modeling*, 443-466. Alexandria, VA: MORS.

Hartley, D. S., III. 1995. "A Revisionist View of Training Simulation Validity Requirements," *Phalanx*, Vol. 28, No. 1.

Hartley, D. S., III and R. L. Helmbold. 1995. "Validating Lanchester's Square Law and Other Attrition Models." *Warfare Modeling*, 467-492. Alexandria, VA: MORS.

Hartley, D. S., III, K. Kruse, J. Denson, G. Austin, and R. Breuer. 1989. *Sensitivity Analysis of the Joint Theater Level Simulation: Part 1*, K/DSRD-70. Oak Ridge, TN: Martin Marietta Energy Systems, Inc.

Hartley, D. S., III, K. Kruse, J. Martellaro, S. Packard, B. Thomas, and V. Turley. 1994. *An Independent Verification and Validation of the Future Theater Level Model Conceptual Model*, K/DSRD-1637. Oak Ridge, TN: Martin Marietta Energy Systems.

Hartley, D. S., III, J. D. Quillinan and K. L. Kruse. 1990. *Verification and Validation of SIMNET-T*, K/DSRD-117. Oak Ridge, TN: Martin Marietta Energy Systems, Inc.

Hartley, D. S., III, Charles Radford and Cathrine E. Snyder. 1991. *Evaluation of the Advanced Battle Simulation in WAREX 3-90*, K/DSRD-597. Oak Ridge, TN: Martin Marietta Energy Systems, Inc.

Hartley, D. S., III and Howard Whitley. 1996. "Verification, Validation, and Certification (VV&C)." *Simulation Data and Its Management Mini-Symposium (SIMDATAM '95)*, ed. Julian I. Palmore, 11-18. Alexandria, VA: MORS.

Helmbold, Robert L. 1995. "Rates of Advance in Land Combat Operations." *Warfare Modeling*, 493-527. Alexandria, VA: MORS.

Henderson, Dale. 1992. "The Multidimensional Space of Validation." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, IV 17-20. Alexandria, VA: MORS.

Hoeber, Francis P. 1981. *Military Applications of Modeling: Selected Case Studies*. New York: Gordon and Breach Science Publishers.

Kapper, Francis B. 1981. "Wargaming I: The Simulation of Crisis." *Defense/81*, May 1981, 14-21. Arlington, VA: American Forces Information Service.

Kostelski, Drazen, John Buzacott, Kenneth McKay, and Xiao-Gao Liu. 1987. "Development and Validation of a Systems Macro Model Using Isolated Micro Models." *Proceedings of the 1987 Winter Simulation Conference*, ed. A. Thesen, H. Grant and W. D. Kelton, 669-676. Atlanta, GA.

Law, Averill M. and W. David Kelton. 1982. *Simulation Modeling and Analysis*. New York: McGraw-Hill Book Company.

Lewis, Robert O. 1997. "Designing VV&A to Cost," *Phalanx*, Vol. 30, No. 2.

Metzger, Jim. 1992. "Verification." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, III 1-12. Alexandria, VA: MORS.

Muessig, Paul R. 1997. "Surprise, Surprise! VV&A Yields a Few Unexpected Results," *Phalanx*, Vol. 30, No. 2.

Naylor, Thomas H. and J. M. Finger. 1967. "Verification of

Computer Simulation Models." *Management Science* 14, 2 (Oct 1967): B92-B106.

Sargent, Robert G. 1996. "Verifying and Validating Simulation Models." *Proceedings of the 1996 Winter Simulation Conference*, ed. J. Charnes, D. Morrice, D. Brunner, and J. Swain, 55-64. Coronado, CA.

Schrank, William E. and Charles C. Holt. 1967. "Critique of 'Verification of Computer Simulation Models.'" *Management Science* 14, 2 (Oct 1967): B104-B106.

Seglie, Ernest and Patricia Sanders. 1992. "Accreditation." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, V 1-12. Alexandria, VA: MORS.

Sikora, Jim and Marion Williams. 1997. "Verification, Validation and Accreditation - What Is It Today?" *Phalanx*, Vol. 30, No. 2.

Stanley, Walter L. 1997. "The Verification, Validation, and Certification of Data in Modeling and Simulation Applications." *Phalanx*, Vol. 30, No. 2.

Tufarolo, J. 1986. *Component Software Test & Evaluation Report for the Modern Aids to Planning Program (ORNL/Sub/85-37710C/036V2, MAPP) JTLS Model Version 1.5B*. Washington, DC: SYSCON.

Whitley, Howard. 1991. "Ardennes Campaign Simulation (ARCAS)." Presented at 59th MORSS, June 1991, West Point, NY. West Point, NY.

Whitner, Richard B. and Osman Balci. 1989. "Guidelines for Selecting and Using Simulation Model Verification Techniques." *Proceedings of the 1989 Winter Simulation Conference*, ed. E. MacNair, K. Musselman and P. Heidelberger, 559-568. Washington, DC.

Williams, Marion L. and James J. Sikora. 1992. "Overview." *Simulation Validation Workshop Proceedings (SIMVAL II)*, ed. Adelia E Ritchie, I 1-6. Alexandria, VA: MORS.

Youngblood, Simone M. 1997. "The Joint Warfare System (JWARS) VV&A Effort: Challenges of a Joint, Object Oriented Development." *Phalanx*, Vol. 30, No. 2.

AUTHOR BIOGRAPHY

DEAN S. HARTLEY III is a Senior Member of the Research Staff at the Oak Ridge National Laboratory and other Department of Energy facilities, where he is Senior Scientist of the Center for Modeling, Simulation, and Gaming. He received his Ph.D. in Mathematics from the University of Georgia. Hartley is a member of the Board of Directors of the Military Operations Research Society (MORS). He is the Past President of the Military Applications Society (MAS) and a member of the College on Simulation, both subdivisions of the Institute for Operations Research and Management Sciences (INFORMS). He is also a member of the INFORMS Board of Directors. Dr. Hartley's research interests include analysis of historical military combat data, verification and validation of military models, analytical support to military operations other than war (OOTW), and modeling information presentation.