Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

VERIFICATION, VALIDATION & ACCREDITATION: DISCIPLINES IN DIALOGUE OR CAN WE LEARN FROM THE EXPERIENCES OF OTHERS?

Panel Presentation

Chair
James D. Arthur

Co-Chair
Richard E. Nance

Systems Research Center
and
Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061-0251, USA

Respondents

Robert G. Sargent

College of Engineering & Computer Science
Syracuse University
Syracuse, NY 13244, USA

Dolores R. Wallace

National Institute of Standards and Technology
NIST North, Bldg. 820, Rm 517
Gaithersburg, MD 20899, USA

Linda H. Rosenberg

Unisys/GSFC
Bldg. 6, Code 300.1
Greenbelt, MD 20771, USA

Paul R. Muessig

Naval Air Warfare Center, Weapons Division
Code 418000D (Attn: Dr. Paul R. Muessig)
1 Administration Circle
China Lake, CA 93555, USA

ABSTRACT

Relative to Verification, Validation, and Accreditation (VV&A) the Modeling and Simulation and Software Engineering communities share similar goals and impediments to achieving those goals. The focus of this panel is to explore how each community addresses the critical issues underlying VV&A. In this paper we provide nine (9) questions and four (4) sets of responses to those questions. The questions are intended to help reveal differences in VV&A emphasis and motivation between the two communities, and to establish a basis for the exchange of mutually beneficial ideas.

1 INTRODUCTION

Within the Modeling and Simulation (M&S) and Software Engineering (SE) communities the ever-expanding complexity of today's systems is dictating the need for a controlled, structured development process that embraces activities focused on: (a) reducing modeling and coding errors, (b) meeting cost and schedule, and (c) producing a quality product. Verification and validation activities can play a prominent role in achieving these laudable objectives. Accreditation, on the other hand, helps instill confidence that a quality product has been produced.

Because both the M&S and SE communities share similar goals and impediments to achieving those goals,

much can be gained in sharing experiences, perceptions, practices and lessons learned as they relate to Validation, Verification and Accreditation (VV&A). Hence, the principal objective and focus of this panel session is to examine how each community incorporates VV&A in its associated activities. In doing so, we expect to raise an awareness of both the strengths and weaknesses of existing VV&A paradigms, and hopefully, recognize additional development activities and techniques that further our common goals.

2. POSED QUESTIONS

In support of the objectives and expectations of the panel session, each panel member is to provide a one page response that addresses as many of the following questions as possible. Each response should reflect the "prevailing wisdom" espoused by his/her discipline. Elaboration is to be provided where needed.

Question 1

Briefly, how would you define, or distinguish among verification, validation and accreditation?

Question 2

What do you perceive to be the goals and objectives of each of verification, validation and accreditation?

Question 3

Relative to achieving the three objectives stated in the first paragraph, how would you rate the *perceived* importance of (a) verification, (b) validation, and (c) accreditation.

Respond to the above question using a 1-5 rating where: 1 => little importance, 3 => important, 5 => essential.

Question 4

Based on what is being *practiced* in your discipline, to what extent are (a) verification, and (b) validation techniques being applied during the development process?

Respond to the above question using a 1-5 rating where: 1 => very little, 3 => often, 5 => extensively.

Question 5

What is the status of accreditation in your discipline?

Points to consider in answering the above question include (but are not limited to): (a) how highly is accreditation touted as a desirable characteristic of a software development firm, and (b) how much of a role does accreditation play in the selection of a software vendor.

Question 6

Testing is a critical activity that supports the evolution of a product that meets stated requirements.

In the development life-cycle:

(a) When should testing begin?
(b) How are test criteria derived?
(c) What are the criteria for judging the test outcome, e.g., completeness and correctness?

Question 7

The extent to which verification and validation (V&V) is practiced varies within and among disciplines.

Within your discipline, what do you see as the top three (3) impediments to implementing V&V practices. List them in descending order of impact.

Question 8

As a matter of fact, constraints are often placed on a development effort that restricts the extent to which V&V can emphasized and employed.

As a project manager, if you had to choose to emphasize verification over validation or validation over verification:

(a) Which would you choose and why?
(b) What would be the impact of that decision?

Consider, for example, the impact on the development process and on the product itself.

Question 9

In performing V&V activities, artifacts can (a) play a supportive role in that they are used to help carry out or guide V&V activities, (b) be the object of scrutiny, or (c) participate in both capacities.

In terms of the above characterization, identify the role that each of the following play: Documentation, Models, Programs, and Data.

2 PANEL RESPONSES

2.1 Response of Robert G. Sargent (Simulation Perspective)

Question 1

The definitions that I use for verification and validation in my work (see, e.g., "Verifying and Validating Simulation Models" in the *Proceedings of the 1996 Winter Simulation Conference*) are "model verification is ensuring that the computer program of the computerized model and its implementation are correct" and "model validation is substantiating that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model." The definition that I propose for accreditation is that "model accreditation is determining that a computerized model satisfies a specified model accreditation criteria according to a specified process for determining that a model satisfies the specified model accreditation criteria."

Question 2

The overall goals and objectives of verification and validation are to obtain high confidence that a "correct" model has been developed, and of accreditation are (a) to convey that a model has been "certified" by a specified accreditation procedure and (b) to aid in the development of confidence of (potential) users of a model that a model is "correct" if it has been accredited. (Note: One can prove a model is incorrect but one cannot prove a model is "correct".)

Question 3

Verification and validation are essential (rating of 5) for (a) - reducing modeling and coding errors and (c) - producing a quality product; doing these take time and money and thus they effect (b). Accreditation plays at best a minor role in reducing modeling and coding errors and in obtaining a quality product; therefore, I give accreditation a ranking of 1 for (a) and (c). Of course, if accreditation is performed, the schedule and cost (i.e., (b)) will be impacted.

Question 4

I believe (a) verification receives much attention in model development and thus I give it a rating of 4.0 and (b) validation receives less attention then verification and thus I give it a rating of 3.0.

Question 5

Model accreditation has started to receive attention from the Department of Defense, but I am not aware of accreditation being used elsewhere.

Question 6

I believe that testing (i.e., performing verification and validation) should take place during the entire model development process as discussed in "Verifying and Validating Simulation Models."

Question 7

I believe the three primary impediments to verification and validation (V&V) are (a) the lack of emphasis and support for V&V by many model sponsors, (b) the lack of interest in performing validation (and to some extent for verification) by many model developers, and (c) the lack of interest in and emphasis for validation (and to some extend for verification) by many educators.

Question 8

I believe that *both* verification and validation are needed for developing models. It is not a choice of one or the other. Either both are done or the model should not be developed.

Question 9

I believe that Data, Models, (computer) Programs, and Documentation are all important in verification and validation. Many individuals find the "Simplified Version of the Modeling Process" that I use in my work (see, e.g., "Verifying and Validating Simulation Models") to visually relate data, problem entity, conceptual model, and computerized model to steps of the modeling process, including verification and validation, to be extremely helpful.

2.2 Response of Dolores R. Wallace (Software Engineering Perspective)

Question 1

Verification consists of those activities to examine statically all software artifacts to see that they meet requirements established during previous activities, as they evolve into the software system. Validation consists of those activities that execute the software system, or some parts of it, to determine if it satisfies

system requirements and solves the right problem. Accreditation consists of those activities to provide measurements that the software system meets predefined criteria, derived from results of other quality assurance and V&V activities (and evidence from approved development process).and assigns some warranty "status" to the system.

Question 2

See Q1. V&V activities facilitate detection & correction of software errors; enhance management insight into process and product risk. Accreditation produce evidences, based on results of activities such as V&V, that appropriate processes have been applied, with resulting measures on the product, to show the product meets pre-defined criteria.

Question 3&8

Rate 5 for all. The answer is "relative," that is, relative to cost of developing the system, the quality requirements of the system, need for an implied or stated warranty of the system, maturity of development methods and tools, developer capability. Studies show that it is much cheaper to find faults early in development, before testing begins. Studies also indicate that if found too late, some design flaws may never be fixed, hence impacting quality. Reasonable combinations of verification and validation strategies may focus on different fault classes. A project must address the importance of these activities in tradeoffs, strategic planning and recognize that both are ALWAYS necessary for high integrity software but the degree of what and how may differ. Accreditation relies on the data from these & other quality activities to calculate the metrics that provide insights to the system quality. The developers (and IV&V) perform V&V activities. The accreditation activities may be performed the "owner" of the software or someone appointed by the owner, or some authorized authority.

Question 4

Based on discussions with others: Verification: 1; Validation 5 (varying degrees but always); Accreditation: 1.

Question 5

A guess is that a) is considered important but b) outside Defense Department, few do it.

Question 6

The test process begins when the project begins, with planning for each category of test (unit, integration, system), especially for the resources, such as any simulation facility, other hardware, CM. Test criteria are derived from the requirements, acceptance criteria, and the potential severity of faults & the probability of their occurrence.

Question 7

3) Changing requirements and supporting materials 2) Lack of understanding of product, to establish verification, test criteria. 1) Lack of management support (requirements to perform, understanding of need, resources availability)

Question 8&3

Risks and product characteristics such as consequence of failure, cost, user applications, deadlines; maturity of the application; maturity and tool support for the methods of each V: all influence the emphasis and the choice influences development pace, amount of changes later, etc.

Question 9

Documentation of requirements is vital to establishing test criteria, designs, providing traceability to implementation and tests. Later artifacts are needed for traceability, design of some analyses and tests, and mapping to models. Models enable V&Vers to verify that developers understand the problem, that V&Vers and developers understand each other; that solution is correct. Programs are essential for any static analyses, for verifying appropriate configuration. Data enables definitions of test parameters, and provides input for test cases.

Summary

Application of V&V requires a comprehensive examination of system requirements & project environment to develop an integrated set of activities to achieve evidence of system quality.

2.3 Response of Linda H. Rosenberg (Software Engineering Perspective)

My environment is the Software Quality Assurance at Goddard Space Flight Center (GSFC) NASA. The

SATC is charged with making measurable improvements in the quality and reliability of software developed for GSFC and NASA. We do not develop software nor do quality assurance activities although we are part of the Quality Assurance Division. We assist projects in implementing standards, guidebooks, and metrics programs to measure the quality of the products and project and monitor potential risks.

Question 1

In the NASA Software Assurance Guidebook, validation is defined as ìthe process of ensuring that software being developed or charged will satisfy functional and other requirements.î Verification is ensuring that ìeach step in the process of building the software yields the right products.î Accreditation is not defined in the guidebook, but I would define it as the verification of the necessary skills and training to complete the tasks.

Question 2

The goals of validation are to make sure the requirements are satisfied; i.e., if a satellite must transmit data at a given time and rate, that the software meets those specifications. A developer does not have to understand the environment of the application to validate the software. Verification requires extensive application knowledge to determine if the software meets the necessary specifications, regardless of ambiguities or inaccuracies in the requirements. Verification would identify if the specified transmission time and rate in the previous example were feasible for this application. The goals of accreditation are to warranty the qualifications of the personnel developing the software.

Question 3

Perceived importance within my discipline:

Validation	3
Verification	3
Accreditation	1

Question 4

Extent applied during development process:

Validation	1-2
Verification	1-2

Question 5

Accreditation as applied to contractors and software vendors is very desirable but unfortunately usually cannot be enforced due to legal problems. Most NASA software is developed by contractors and NASA cannot tell them who to hire. Some accreditation is attempted with contract awards but usually is very general and ineffective.

Question 6

At NASA, formal testing begins after Unit Testing, after the Test Readiness Review (TRR). Test plans/criteria are developed starting in the requirements phase. The criteria for judging the test outcomes should have repeatable input/outputs and have clear pass/fail criteria. Test program results must validate the software as an acceptable satisfaction of the requirements.

Question 7

Unfortunately, V&V is rarely done at GSFC except in large, high profile projects. Although NASA has an IV&V Facility in WVA, itís services often go unused. Two primary impediments for V&V are the costs and administration of the V&V activities, and contractual/legal issues. A secondary problem is that the value from V&V is perceived to be very low and not cost effective.

Question 8

If, as a project manager, I had to choose between validation and verification, I would choose verification, since the propagation of errors compounds costs and delays, and severely impacts performance. However, from a funding perspective, validation provides the largest return on investment since the earlier problems are found, the less the cost to fix, hence, if errors in the high level requirement documents are identified through early validation, there is greater savings.

Question 9

The primary artifacts for V&V are the documentation and the data; these provide the most information on what is going on. V&V on models has been tried but was not very effective. Source code is rarely, if ever used for V&V activities.

2.4 Response of Paul R. Muessig (Simulation Perspective)

The wording of the questions provided for comment, and the forwarded requests for clarification regarding them, are indicative (to me, at least) of a phenomenon most people in the M&S and SE communities recognize, but about which no one has done anything constructive that I

am aware. The phenomenon of which I speak is the disconnect between the definitions of verification, validation and accreditation as used in the M&S and SE communities. This disconnect will color my responses to several of the questions relative to other respondents.

Questions 1 and 2

For example, to answer Questions 1 and 2, I would refer to the definitions of VV&A worked out by the Military Operations Research Society (MORS) for use in military modeling and simulation. Verification is the process of determining that a model implementation accurately represents the developer's conceptual description and specifications. Validation is 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 is the official certification that a model, simulation or federation is acceptable for use for a specific purpose.

In my experience, the M&S community's definition of verification seems, in practice, equivalent to the SE community's definitions of "V&V" (or "IV&V"). The M&S community's definition of validation includes an explicit reference to the "perspective of the intended use of the model". This comparison with the real world (validation testing), so important to military M&S, is only one aspect of SE validation, however, which involves a much broader range of potential activities. In the M&S community, accreditation does not refer to any characteristic of a software developer, nor is it a warranty that the software has been built to specifications and won't "break", but rather that the model's strengths and weaknesses have been evaluated in light of its intended use, and has been judged "good enough" for that use. In the military M&S community, at least, accreditation for one purpose is generally not sufficient to accredit a model for another (unrelated) purpose without further justification (i.e., another accreditation).

Question 3

In answering Question 3, again my answer would depend on which community's definition of "V&V" we are talking about. My understanding of verification from the SE perspective is that it answers the question, "Did I build the thing right?", while validation answers the question, "Did I build the right thing?" From this standpoint, both verification and validation are equally necessary, and might each rate a 3 or a 4 on your scale. (Not being familiar with the SE understanding of accreditation, I cannot usefully comment on where I would rank it on your scale.) From the M&S perspective, however, it has been my experience that verification activities are deemed less important than validation activities when evaluating M&S credibility (keeping in mind the differing definitions and perspectives cited above), and I would rate verification a 2 or a 3, and validation a 4. In the military M&S community, accreditation is a highly charged issue, with overtly political overtones. From a purely technical standpoint, the way accreditation is done now by most people in the M&S community is woefully inadequate, and I would rate it a 1 or a 2 in technical importance and impact. If it were done correctly (i.e., without so much political influence), I would rate it a 3 or a 4 in importance and impact. From a political standpoint, however, accreditation is the "Holy Grail"; everyone needs to have their model accreditation (DoDD 5000.59, DoDI 5000.61 and a host of Service-specific policies, instructions and directives. From this standpoint, accreditation rates a 5. Definitions are everything.

Question 4

In response to Question 4, the answer depends on whether the model being considered is under development or "legacy" (software developed before detailed SE standards and specifications were codified). In the former case, the answer is 5: there is extensive application of SE-type V&V activities to new M&S development. Whether or not the "right" SE V&V is being done is a hot issue in the M&S V&V community, particularly as it relates to accreditation. In the case of legacy M&S, I would answer 2: little structured V&V (of any type) is being conducted.

Question 5

Question 5 proves my point about the differences in the definition of key terms between the SE and M&S communities. In the M&S community, accreditation has nothing to do with the characteristics of a software vendor. Accreditation is an official determination that a model is "good enough for government work", based on a clear definition of how the model will be used, and an objective analysis of how credible the model is relative to these usage requirements. From this latter standpoint, as noted above, accreditation is very important on the surface of things. From an implementation standpoint, little substantive attention is paid to objective requirements for accreditation. It is very frequently decided more by politics than by objective determinations of model suitability.

Question 6

In answering Question 6, I would argue that, depending on the development paradigm being used, testing should begin as early in the development life-cycle as possible. Test criteria should be derived and prioritized from a specification of the requirements for accuracy and credibility based on how the model will be used. If the expected application range of the finished model is broad and will have great impact on human safety or acquisition decisions, for example, more should be spent on testing to ensure coding accuracy. If the model has less stringent credibility requirements, the scope of testing can be narrowed in favor of other, perhaps less costly, approaches to establishing software accuracy.

Question 7

For Question 7, the prioritized impediments to implementing V&V practices in the M&S community are 1) lack of understanding of the value of V&V, 2) funding, and 3) funding.

Question 8

For Question 8, from the M&S standpoint my experience has shown that verification activities add just as much credibility to a model as validation activities do (remember the M&S community definition of validation). Since verification costs less than validation, the marginal utility of each verification dollar spent is greater than the marginal utility of each validation dollar spent. Consequently, I would focus my credibility assessment activities on verification (remember, M&S "verification" is equivalent to SE "V&V"). The impact of this decision would be to reduce the amount of that type of credibility so sought after in the M&S community: comparison of model predictions with real world data (M&S "validation"). Consequently, I would have to consider more seriously a user's inputs to the process of defining credibility requirements.

Question 9

I'm not sure what you mean by Question 9, although it appears to be worded from an SE standpoint. I don't think I can contribute substantively to this question without a better appreciation of the meaning of the terms.

I am guessing that the perspectives I have included as part of my answer to these questions will come as a surprise to those whose background is primarily in the SE community. A software engineer tends to looks at V&V primarily as a quality assurance process focused on the end product: the software. The M&S community tends to look at V&V as a means to assess the credibility and suitability of a model for use in particular application. The reason for the difference in emphasis is that all models are software, but not all software is a model. The SE process applies as well to the development of word processing software as to the development of a missile flyout model. But the requirement that the latter meet objective criteria related to comparison of software outputs with real world data puts the focus of validation activities in a different realm. Seen from this perspective, SE V&V is a subset of (or contributor to) M&S V&V. The difference lies in scope and application of the data generated by V&V in the two cases. Getting the M&S and SE communities to understand each other in this important area will be an important aspect of the panel discussion in December.

AUTHORS BIOGRAPHIES

ROBERT G. SARGENT is a Professor in the L. C. Smith College of Engineering and Computer Science at Syracuse University. He received his education at the university of Michigan. Dr. Sargent has served his profession in numerous ways and has been awarded the TIMS (now INFORMS) College on Simulation Distinguished Service Award for long-standing exceptional service to the simulation community. His research interests include the methodology areas of modeling and discrete event simulation, model validation, and performance evaluation. Professor Sargent is listed in *Who's Who in America*.

DOLORES R. WALLACE leads the Software Standard Reference Materials and Data Project which provides metrology and reference data for software assurance; she researches and develops methods and tools that improve software quality and correctness. She served as Program Manager, Center for High Integrity Software System Assurance (CHISSA), which was a collaborative effort for research and technology transfer among industry, government and academia; is the recipient of the 1994 Department of Commerce Bronze Medal Award. She has conducted research and written guidelines and standards in high integrity software, software verification, validation, and testing, software management, software quality assurance; she has published articles in the Encyclopedia of Software Engineering (Wiley) and in the IEEE Tutorials on Software Requirements Engineering and Software Engineering. She is co-author of the book, Software Quality Control, Error Analysis, and Testing, Noyes Data Corporation, 1995. She has served as General Chair of COMPASS Board of Directors; lecturer in the

IEEE Seminars on Software Verification and Validation, and as Co-Guest Editor, IEEE Software, May 1989 Issue on Software V&V. She has an MS in Mathematics, Case Western Reserve University.

LINDA H. ROSENBERG is an Engineering Section Head at Unisys Government Systems in Lanham, MD. She is contracted to manage the Software Assurance Technology Center (SATC) through the System Reliability and Safety Office in the Flight Assurance Division at Goddard Space Flight Center, NASA, in Greenbelt, MD. The SATC has four primary responsibilities: Metrics, Standards and Guidance, IV&V, and Outreach programs. Although she oversees all work areas of the SATC, Dr. Rosenberg's area of expertise is metrics. She is responsible for overseeing metric programs to establish a basis for numerical guidelines and standards for software developed at NASA, to investigate the role of metrics in risk assessment and management of software projects, and to work with project managers to use metrics in the evaluation of the quality of their software. As part of the SATC outreach program, she has presented metrics/quality assurance tutorials at GSFC, and IEEE and ACM conferences. Dr. Rosenberg holds a Ph.D. in Computer Science from the University of Maryland, an M.E.S. in Computer Science from Loyola College, and a B.S. in Mathematics from Towson State University. She is a member of Electrical and Electronic Engineers (IEEE), the IEEE Computer Society, the Association for Computing Machinery (ACM) and Upsilon Pi Epsilon.

PAUL R. MUESSIG is currently Director of the Joint Accreditation Support Activity (JASA) at the Naval Air Warfare Center, Weapons Division, China Lake, CA. JASA's goal is to support airborne weapons system acquisition programs (and others) with credible integration of M&S into program objectives. JASA services include M&S application analysis, VV&A requirements definition, and VV&A planning, execution and reporting. (For more information, visit the JASA Home Page at http://www.nawcwpns.navy.mil/~jasa). Dr. Muessig has extensive experience in VV&A methodology development, planning and execution in support of numerous M&S applications in acquisition and testing. He has also worked as a defense analyst at the Center for Naval Analyses in Washington, D.C., contributing to survivability analyses for advanced technology aircraft, and leading model validation efforts for the Advanced Low Altitude Radar Model (ALARM). Dr. Muessig holds a B.S. in Chemistry from St. Joseph's University and a doctorate in Physical Chemistry from Brown University.

JAMES D. ARTHUR is an Associate Professor of Computer Science at Virginia Tech (VPI&SU). He received B.S and M.A. degrees in Mathematics from the University of North Carolina at Greensboro in 1972 and 1973, and M.S. and Ph.D. degrees in Computer Science from Purdue University in 1981 and 1983. His research interests include Software Engineering (Methods and Methodologies supporting Software Quality Assessment and IV&V Processes), Parallel Computation, and User Support Environments. Dr. Arthur is the author of over 30 papers on software engineering, software quality assessment, IV&V, and user/machine interaction. He has served as: participating member of IEEE Working Group on Reference Models for V&V Methods; Chair of Education Panel for National Software Council Workshop; Guest Editor for *Annals of Software Engineering* special volume on Process and Product Quality Measurement; and Principal Investigator or Investigator on 14 externally funded research projects totaling in excess of $1.4 million. Dr. Arthur is a member of Pi Mu Epsilon (Math Honor Society), Upsilon Pi Epsilon (Computer Science Honor Society), Golden Key National Honor Society, Sigma Xi (National Research Society), ACM, and the IEEE Computer Society.

RICHARD E. NANCE Biography found elsewhere in these preceedings.