MITRE TECHNICAL REPORT

When Systems are Simulations – T&E, VV&A, or Both?

June 1998

Michael Borowski, 26571
Priscilla Glasow, 25988

It has been approved for public release.

MITRE

Washington C3 Center
McLean, Virginia

Abstract

This paper provides a brief introduction on the use of modeling and simulation (M&S) and the importance of the credibility of that M&S when it is used to support system acquisition. A previous paper, *The Relationship of VV&A to T&E* (Allen *et al*. 1997), is briefly reviewed, followed by an introduction of a new example of M&S use. This example cites the recent development of systems that are themselves simulations. As new acquisition systems, they require traditional test and evaluation (T&E). As simulations, they require verification, validation and accreditation (VV&A) to ensure credibility. The main text of this paper examines key commonalties between these two processes. It provides insights into how T&E and VV&A requirements' can be met with reduced cost and risk, to the optimal benefit of the user. A case study is provided and conclusions presented.

KEYWORDS: Test and Evaluation, Verification and Validation, Modeling and Simulation.

Section 1

Introduction

The answer is *a little of both*.

Risk reduction is the primary purpose for both test and evaluation (T&E) and verification, validation, and accreditation (VV&A). By evaluating system performance against stated requirements, the user can gain confidence in the system produced. Modeling and simulation (M&S)s is being used as a key tool in system acquisition to reduce the time to field a system, resources to develop and evaluate that system, and decision risk. The use of M&S can also help evaluate and improve the quality, military utility, and supportability of fielded systems. In the T&E phase of system acquisition, M&S is used to develop parameters for mission rehearsal, design tests, analyze data collected during testing, and evaluate regions of the operational envelope that are otherwise not testable. While M&S is a useful tool for predicting, training, and planning, it is *not* a substitute for testing. M&S is only useful if it applies to the evaluation of the system being acquired, and if it is capable of replicating reality to an acceptable level as required for the particular use. The evaluation of M&S systems against the requirements, both system-specific and in terms of its real world representation provides insight into M&S credibility.

This paper is not intended to address the specifics for performing VV&A. Instead, it will quickly review earlier work done to characterize the use of M&S in system acquisition and propose a new case of current application.

Section 2

Previous Research

The Relationship of VV&A to T&E identified four cases where M&S has traditionally been used to support system acquisition. That research illustrated that a clear overlap exists between the two processes and suggested areas where collaboration might reduce cost and risk. The dialogue created by that paper has served to promote cooperation between the testing and VV&A communities.

Table 2-1 illustrates the four cases:

Table 2-1. Relationship of VV&A plans to TEMPs

CASE	M&S	OPERATIONAL SYSTEM	RELATIONSHIP (VVAP to TEMP)
1 No Acquisition	- Used for readiness, force structure, or sustainability - VV&A Plan	- No operational system developed - No TEMP	V
2 Precedes Develop-ment	- Used for concept def. of operational system - VV&A Plan	- Normal acquisition - TEMP	T V
3 Supports Develop-ment	- Supports concept definition - Model updated during development and test - VV&A Plan	- Acq. supported & guided by M&S for perf. modeling and engineering trades - TEMP indirectly influenced by VVAP	V T
4 Part of Develop-ment	- M&S embedded in and developed as component(s) of operational system - VV&A Plan	- Normal acquisition - VVAP becomes part of TEMP effort - VV&A and DT&E/ OT&E tests directly support each other	V T

In Case 1, M&S are built for reasons not related to system acquisition. Since there is no system being acquired and no T&E activity, there is no defined relationship between T&E and VV&A.

In Case 2, M&S is developed to support the Concept Exploration and Program Definition phases of acquisition. M&S precede's system development, but is not updated as the system matures. The model loses congruence with the system being developed except for requirements, any VV&A conducted will have little relevance to T&E of the mature system.

In Case 3, M&S supports a system under development. The digital representation of the system precedes system development and is updated as the system matures. In this case, the real system and the model are distinct entities. The VV&A of the model and the T&E of the system occur in parallel. Following the Model-Test-Model paradigm, the T&E and VV&A processes, complement and support each other. The model is used to guide the system development and the developing system's test results are used to refine the model. The developing Simulation Based Acquisition concept takes its roots in using modeling and simulation as the basis for development of new systems. It clearly follows a Case 3 scenario for VV&A and T&E comparison.
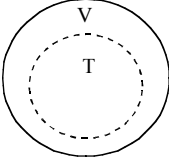
In Case 4, the model is a subset of the system. M&S are totally embedded within the operational system. This integration of the VV&A and T&E processes yields three key benefits: commonality and reuse of testing techniques, value of conceptual modeling and early correction of system problems. The reader is encouraged to read *The Relationship of VV&A to T&E* for a full discussion of each of these benefits.

Section 3

A New Example

Case 5 is the addition to that earlier work. Here, the system under test is itself a simulation. The system hardware consists solely of the computer platform(s) required to run the simulation. The system software consists of only the simulation. Table 3-1 illustrates Case 5:
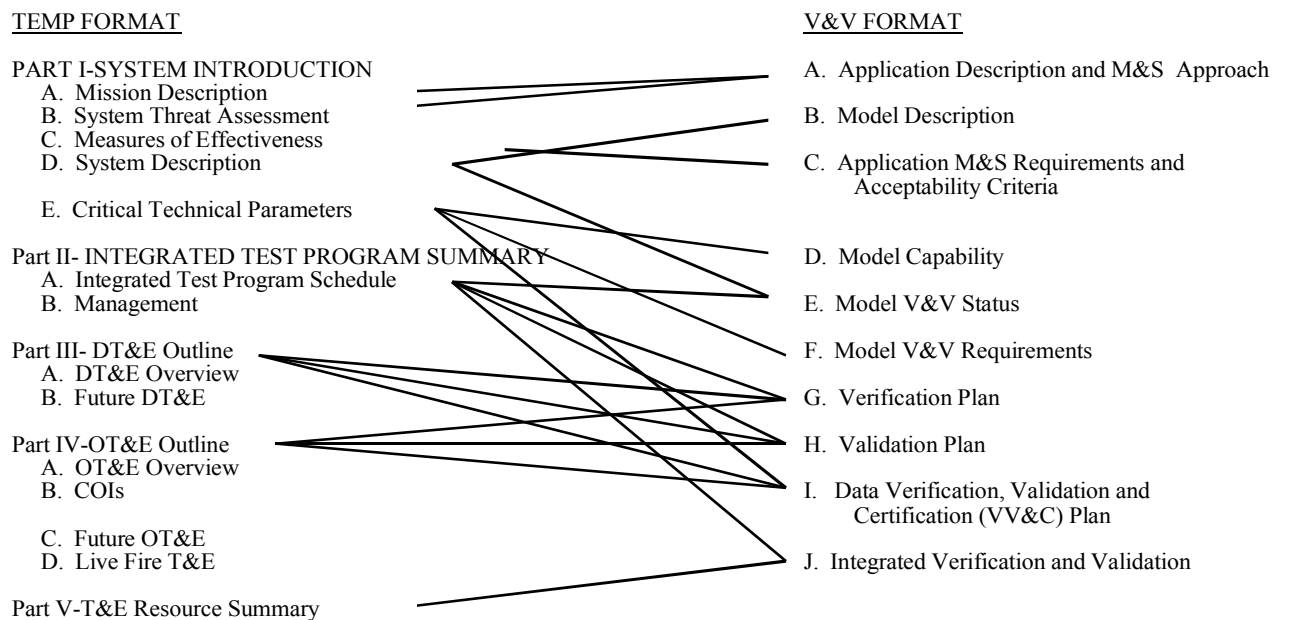
Table 3-1. Acquisition of a Simulation

CASE	M&S	OPERATIONAL SYSTEM	RELATIONSHIP (VVAP to TEMP)
5	M&S is the System	-Acquisition of M&S -VV&A and T&E congruent	V T

The relationship of the T&E and VV&A processes is illustrated as being roughly congruent, with T&E a subset of VV&A. A crosswalk of the two processes was conducted as part of the earlier research and a comparison was made of the information required to support each process. The VV&A Plan format contained in the DoD VV&A Recommended Practices

Guide (RPG) (Glasow, ed. 1996) was compared to the Test and Evaluation Master Plan (TEMP) format. It was found that the information requirements were virtually identical, but that the VV&A process includes certain activities, such as code verification or algorithm validation, that are not part of the T&E process. However, where problems are identified during T&E, there may be a need to examine the code or algorithms, hence T&E is an open subset of the encompassing VV&A process. Table 3-2 illustrates this crosswalk:

Table 3-2. TEMP VV&A Crosswalk

TEMP FORMAT

PART I-SYSTEM INTRODUCTION
    A. Mission Description
    B. System Threat Assessment
    C. Measures of Effectiveness
    D. System Description

    E. Critical Technical Parameters

Part II- INTEGRATED TEST PROGRAM SUMMARY
    A. Integrated Test Program Schedule
    B. Management

Part III- DT&E Outline
    A. DT&E Overview
    B. Future DT&E

Part IV-OT&E Outline
    A. OT&E Overview
    B. COIs

    C. Future OT&E
    D. Live Fire T&E

Part V-T&E Resource Summary

V&V FORMAT

A. Application Description and M&S Approach

B. Model Description

C. Application M&S Requirements and
    Acceptability Criteria

D. Model Capability

E. Model V&V Status

F. Model V&V Requirements

G. Verification Plan

H. Validation Plan

I. Data Verification, Validation and
    Certification (VV&C) Plan

J. Integrated Verification and Validation

Section 4

Key Commonalties

Requirements

The T&E process is founded in requirements, including Critical Technical Parameters, Critical Operational Issues, and Measures of Performance and Effectiveness. The maturity of this process provides an excellent benchmark for the evolution of the VV&A process. In the same way that the T&E process assesses operational system performance, the VV&A process assesses M&S credibility. The DOD Generic VV&A Process, described in the RPG, begins by identifying the problem to be solved and the requirements for solving that problem. The next step is to determine the problem solving approach. M&S is one tool for problem solving, but other tools may also be used to arrive at a solution. Given that at least part of the solution will be arrived at through M&S, general requirements for model capabilities are identified. Depending on these general requirements, the problem solver may be able to use an existing model either "as is" or modified, or a new model may need to be developed. Once that decision is made, requirements for the specific model(s) chosen are established and the model is prepared for use.

While the T&E process is rooted in requirements' definition, the VV&A process has not yet learned the lesson or importance of requirements. Many programs attempt to avoid requirements' definition or make unfounded assumptions. One assumption is that M&S is the correct tool to use where another tool might be easier or less costly for the given problem. Another assumption is in choosing a specific model without rationale or basis in requirements. A poor choice of model may result in invalid results where the model chosen was not built to answer particular types of questions. Sometimes these problems are due to unfamiliarity with the VV&A process, although there have also been instances where sub-optimal decisions were intentionally made. Such decisions often reflected a desire to maximize resources in other areas or to placate a decision maker who had already decided what tool would be used. In a few cases, requirements have been "tailored" out of the VV&A process. "Tailoring" is a VV&A term that describes the focusing of a well-planned VV&A effort on those *tasks* that will provide optimal return on investment. It is the process of selecting which V&V tasks and techniques will provide the most expedient and credible results by which the model can be assessed. Requirements' definition, however, is not open to negotiation or tailoring. Common sense dictates that in order to credibly assess a simulation, one must know what the simulation is supposed to do!

Management

The T&E process is well established and is understood by a large community of developers, testers, and managers. By comparison, the VV&A process is relatively new. The T&E process utilizes mature methods that provide excellent examples which VV&A would do well to emulate. For example, the TEMP requires that responsibilities for each segment of the testing community be delineated. Another example is the approval process for the TEMP and other testing documents, which requires negotiation and compromise among participating organizations prior to the start of a T&E effort. By comparison, VV&A efforts reflect a wide variety of dissimilar and nonstandard approaches, many of which are incomplete or which unnecessarily delay the start of the VV&A process. Identification of roles and responsibilities -- who will do what for whom, when, where, why and for how much - is *essential* prior to the start of any VV&A effort. Unfortunately, many programs do not learn this lesson until after they have expended large sums of time and money, losing the optimal window of opportunity, during the development of the simulation. Often resources are wasted on educating the contractors who are supposed to understand and implement VV&A, thereby adding to the reputation of VV&A being costly.

Documentation

The T&E process is characterized by clearly defined documentation. Although common reporting formats for VV&A were developed for DOD, many programs avoid committing implementation details to writing. Initial attempts at writing a VV&A plan often include large tutorials that have been written at considerable expense. No new information is offered in such treatises despite claims of "tailoring" to meet "unique" program needs. However, the veracity of such claims have been few.   Where the T&E process details the specific information requirements and criteria for assessing the system under test, most VV&A plans to date have failed to provide the executable detail needed to perform V&V. Specific V&V tasks and techniques must be identified and linked to specific portions of the problem to ensure that those tasks are indeed necessary. Unfortunately, the combined lack of stated requirements and the absence of executable V&V detail result in VV&A plans that merely provide a high level strategy, but never provide clear direction and action.

The evolving Simulation Based Acquisition concept will employ synthetic environments and digital representations of evolving systems. This will require disciplined implementation of requirements' traceability, sound management processes and thorough documentation. With only digital models and databases, the software products will be based upon proven software development processes. The VV&A community will have to actively engage these practices to ensure that VV&A is not "assumed away" under the context of good software development processes, or replaced altogether as an "unnecessary expense" to program offices.

JWARS Case Study

The JWARS program office initiated a V&V effort in October 1997 to support the production version of JWARS. Additionally, the Joint Requirements Oversight Council (JROC) directed that a test and evaluation plan be required as part of the Operational Requirements Document (ORD). T&E planing began in December 1997, using the V&V Plan as its initial point of departure.

JWARS presents a new form of modeling and simulation use in DOD. Where M&S has historically been used to support development of weapons systems and other tangible assets, the JWARS simulation is *itself* the system under test. As such, the system requires some form of T&E, while its formulation as a simulation requires that JWARS also undergo V&V, under DOD Instruction 5000.61. The relationship of these two processes has been the subject of research being conducted to support other DOD programs facing similar dilemmas. That research has been applied to the evolution of an integrated JWARS T&E / V&V strategy.

The JWARS V&V effort is distinct from the in-house quality assurance being conducted by the developer. It is independent in the sense that the V&V contractor reports to the oversight body for JWARS development, the Joint Analytic Modeling Improvement Program (JAMIP). The conduct of additional V&V beyond developer QA is evidence of the commitment on the part of the program office to provide a product that is useful and usable by the warfare analysis community. During the first quarter of FY98, the V&V contractor developed a V&V plan. The RPG was used as the primary resource for the development of the plan. The V&V contractor also worked closely with an oversight group of recognized DOD VV&A experts who provided input and direction for the plan's development. The V&V plan is currently under Service and CINC review. It has received favorable comments from many of the key reviewers, although significant concern for supporting resources and the V&V relationship to T&E has been stressed.

JWARS is an ACAT III program, therefore formal test and evaluation per DOD acquisition directives is not required. However, the JWARS Program Office has elected to use the Test and Evaluation Master Plan (TEMP) format to guide the development of a T&E plan. T&E planning and execution is occurring in parallel with the V&V effort and is leveraging from the V&V Plan to ensure coordination between these two processes.

JWARS T&E differs from traditional program T&E in two significant ways. First, to support its T&E initiative, the JWARS program has recognized the need to involve both the Services' T&E and analytical agencies. The T&E agencies are the traditional sources for test and evaluation support to the Services and would naturally be sought for their expertise during this process. However, the Operational Test Directors (OTDs) at these agencies are primarily warfighters who test hardware -- platforms, weapons systems, and equipment --

which is distinctly different from "analytical software" such as JWARS. Whereas the OTDs represent the military users of hardware systems, the *military analysts* are the targeted user community for JWARS. The operational test of a weapons system often involves hands-on use of the system in the field by military operators. Similarly, operational test of a simulation requires use by personnel trained and experienced in that arena. For JWARS, those personnel are the analysts who are trained in the use of theater-level simulation. The analysis agencies will serve as test sites and the OT&E agencies will provide oversight and report on the testing conducted.

A second difference is JWARS' intended use of alpha and beta testing. Although conducted at field sites and involving potential users, the primary purpose of these tests is to provide feedback to the developer. JWARS intends to use the alpha and beta testing phases to support both the developer's quality assurance program, and to provide the military user community with early opportunities to become familiar with the simulation

The JWARS V&V Plan identifies V&V techniques from the RPG, however due to limited resources, problem domain validation will be restricted to face validation by subject matter experts. This technique, while necessary, is not sufficient for credible validation of a simulation of the magnitude and criticality of JWARS. Therefore, the T&E effort has been focused on *extending the validation envelope* through additional test techniques that meet both V&V and T&E objectives. In their capacity as T&E/ V&V oversight support to the program office, MITRE developed a crosswalk between the V&V Plan (as recommended in RPG) and the TEMP (as described in DODR 5000.2-R). This provided important information regarding the information overlaps which exist between the two documents, and identified where existing information in the V&V Plan could be leveraged as immediate input to the T&E plan. The value of this approach is the reduction of duplication of effort, thereby saving time and money, while ensuring that these processes mesh and complement each other.

A Working Group Integrated Process Team (WGIPT) was established to develop the strategy, identify test activities and testers, ensure the correct conduct and documentation of test events, review test results, and provide recommendations to the JAMIP. The WGIPT consists of representatives from the Services' analysis organizations, the Services' T&E agencies, OSD, J-8, and the JWARS program office. Advisors to the WGIPT included the JWARS developer, MITRE, and representatives from the Joint Data System (JDS). An IT&E contractor will prepare test plans for the WGIPT's approval, provide periodic status briefings to the WGIPT, coordinate required memoranda of agreement, and document T&E results. An initial concept for JWARS T&E was developed and presented to the WGIPT in March 1998. This concept focused on testing of the Planning & Execution and Force Assessment applications of JWARS prior to IOC. Systems Effectiveness & Tradeoff Analysis and Concept & Doctrine Development and Assessment was identified for later testing. A set of proposed performance measures was also provided, with traceability, utility, and V&V highlighted as the three key performance parameters. An additional briefing

described the Fielding Plan developed by J-8, which provides a detailed description of the logistic implementation for JWARS fielding at identified test sites. The Fielding Plan included designation of the level of testing at which each test site would participate. J-8 also provided a macro-level T&E process diagram that enumerated the various tasks and responsibilities for T&E. This was used to develop a strawman partitioning of lower level tasks and apportionment of those tasks among the various players. These products are currently under Service and CINC review.

Both the T&E and V&V processes require identification and agreement of roles and responsibilities before beginning either process. Key to the testing of JWARS are the roles of the traditional T&E agencies and the eventual users of the simulation, the analytical organizations. Similarly, there is a need for a balanced perspective among the Services, and in relation to OSD and the Joint Staff. While the WGIPT reflects a reasonable balance among these various players, there remain significant decisions to be made regarding the tests that each test site will perform and the method for supporting the test events.

Outstanding technical issues include the relationship of the T&E effort with the V&V plan, which is only slightly more mature, but progressing. Despite this early positive direction, the V&V effort did not produce an executable plan of *specific* tasks and techniques. The T&E effort must, therefore, focus on specifying measures of performance, critical technical parameters, and critical operational issues; and developing comprehensive, executable test plans that incorporate these necessary elements.

The JWARS T&E schedule provides timeframes for task identification and assignment, determination of assessment criteria for those tasks, negotiation for test site support, and identification of test events and resources. This schedule is both reasonable and necessary. The use of the T&E community's proven methodology for planning a T&E effort is highly appropriate and is the most effective use of time and resources. The JWARS program office is committed to building a useful, usable simulation product for DoD, and an optimal level of success can be achieved with the participation of JWARS' potential customers, the Services and CINCs.

Conclusions

From the comparisons and the JWARS case study, the parallels between testing and VV&A are of such congruence that the two processes can effectively merge for the acquisition of a simulation system. Short of claiming that the T&E of an acquisition is equivalent to V&V of a simulation, the coordinated efforts of T&E and VV&A can leverage resources to effectively produce the best possible M&S product at lower cost and shortened schedule with fewer fielding risks.

The JWARS case study shows how, under the WGIPT, the T&E and M&S communities can combine their efforts.   Other DOD efforts, like SBA, can surely emulate the JWARS experience and lessons learned, blending the acquisition knowledge and resources of the T&E community with the VV&A expertise of the M&S community.   Although the JWARS effort is not complete, it is envisioned that the final product will prove to be highly credible and produced at a reasonable cost.

List of References

1. Allen, H. P.; P. B. Burleson and P. Glasow. 1997. "The Relationship of VV&A to T&E." *Proceedings of the 1997 Summer Simulation Conference*, Society for Computer Simulation.

2. *Department of Defense Instruction 5000.61*, DOD Modeling and Simulation (M&S) Verification, Validation and Accreditation (VV&A), April 29, 1996.

3. *Department of Defense Regulation 5000.2-R*, Mandatory Procedures for Major Defense Acquisition Programs (MDAPS) and Major Automated Information Systems (MAIS) Acquisition Programs, March 15, 1996.

4. Glasow, P. (ed.). 1996. "*Department of Defense Verification, Validation and Accreditation Recommended Practices Guide*." Defense Modeling and Simulation Office.