

## Session T7: V&V Policies, Guides, Handbooks, and Standards

Session T7 leaders:

Co-Chair:

**Susan Solick** (US Army/TRAC)

**Marcia Stutzman** (Northrop Grumman Information Technology)

Session Recorder: **François Hemez** (Los Alamos National Laboratory)

T7 Materials in Foundations '02 proceedings:

### Papers

*Codes, Standards, Recommended Practices, and Guides of Engineering & Scientific Professional Societies: Application to Verification & Validation in Computational Engineering* (27 pp) [T7\_Schwer.pdf]

**Len Schwer** (Schwere Engineering and Consulting)

*Accrediting Complex M&S for an Analysis of Alternatives: A Successful Approach* (11 pp) [T7\_liptak.pdf]

**Lynda Liptak & Peter Delinski** (USAF Office of Aerospace Studies)

Slides (may contain back-up materials and notes)

*Codes, Standards, Recommended Practices, and Guides of Engineering & Scientific Professional Societies: Application to Verification & Validation in Computational Engineering* (24 slides) [T7B\_1\_Schwer in both pdf and ppt formats]

**Len Schwer** (Schwere Engineering and Consulting)

*V&V Policies, Guides, Handbooks, and Standards* – two versions, the longer one includes material from other session briefings (34 slides) [T7B\_2\_Youngblood in both pdf and ppt formats] and (77 slides) [T7B\_all\_Youngblood in both pdf and ppt formats]

**Simone Youngblood** (DMSO)

*Department of the Navy Verification, Validation, and Accreditation (VV&A) Policy and Standards* (13 slides) [T7B\_3\_Park in both pdf and ppt formats]

**Jennifer Park** (NAVMSMO)

*Department of the Army Verification, Validation, and Accreditation (VV&A) Policy and Standards* (17 slides) [T7B\_4\_Solick in both pdf and ppt formats]

**Susan D. Solick** (TRAC-FLVN)

*Department of the Air Force Verification, Validation, and Accreditation (VV&A) Policy and Standards* (14 slides) [T7B\_5\_Johnson in both pdf and ppt formats]

**Sam Johnson** (AF/XIWM)

*Air Force Example* (7 slides) [T7B\_6\_Liptak in both pdf and ppt formats]

**Lynda Liptak** (AFMC/DR-OAS)

Participants in this session are listed at the end of the Discussion Synopsis.

**Discussion Synopsis** (to provide perspective on papers & briefings identified above).

**Objective:** Provide information on available VV&A policies, standards, and guidance. Discussion focused on identifying voids and recommendations to decision-makers and managers.

**Audience:** Standing room only. Most attendees express their desire to learn about policies and standards. Whether anybody has experience implementing a V&V plan is unsure at this point.

**Synopsis:** Two perspectives were presented: One encapsulating the standards and guidance available in the commercial sector and one presenting DoD policies and guidance. DoE and NASA were not represented. There is a need to identify other communities with VV&A standards and guidance to explore.

**Key insights:**

- There is consistency across DoD and Service / Component policies. Policies are structured in a hierarchical manner so that information flows down from DoD to the Services/Components and then down to subordinate command levels (e.g., COMOPTEVFOR) and finally down to the individual programs (e.g., Tomahawk).
- The terminology used to describe “standards” by DoD and the Services is not the same as in professional societies but the process established to define them at the DoD/Service level emulates and complements the processes used by professional societies.
- Existing best practices and guidance provide a solid foundation and description of the VV&A process; however there is still a need for more detailed descriptions of the application of specific techniques as well as the desire for detailed examples of successful implementations of VV&A as well as lessons learned.
- Within DoD and the Services, there are multiple mechanisms, which facilitate communication and feedback throughout. For example, various forums exist which allow for the identification of issues, exchange of ideas, and sharing of lessons learned (e.g., VV&A Technical Working Group (TWG), Navy TWG). In addition, guidance documents provide automatic feedback loops that allow for community input into the evolution of the documents.
- Relative to the relationship of CMMI and VV&A of M&S, CMMI is a tool for evaluating effectiveness and maturity of an organization’s management of its software development process. CMMI looks at the process and V&V looks at the end product. Where you have an organization with a mature management process, you are more likely to have the artifacts that will support the V&V effort.

## Secondary insights:

- V&V should be an integral part of the M&S development process.
- Legacy M&S often lack necessary documentation and artifacts. The cost to recreate and reengineer these documents to support the VV&A effort shouldn't be attributed to VV&A but in many cases it is, thereby driving the perception that VV&A is overly expensive. Therefore there is a need for a more mature M&S development process as well as contractual guidelines to describe the requisite products.
- A standard description of “conceptual model” is needed to better facilitate its development and the understanding and sharing of its contents by the stakeholders.
- There is a need incentivize PMs to do V&V, to identify sanctions for failures and rewards for success.

**Presentation 1: “Codes, Standards, Recommended Practices and Guides of Engineering and Scientific Professional Societies: Application to Verification and Validation in Computational Engineering,”** Len Schwer (Schwer Engineering and Consulting).

### Standards:

- facilitate communication – serve as common language, defining quality and establishing safety criteria
- generally lower overall costs
- generally increase credibility (consumers accept products more readily when they can be judged on intrinsic merit)

American National Standards Institute (ANSI) approves organizations for becoming standards (only USA representative in ISO). They do not create standards but monitor organizations that do. General approach:

- Standards achieved by consensus of the people involved
- Follow due process – all objections are heard and acted upon
- All open – the larger community being represented can provide input
- Purge process – every standard is resubmitted every 5 years and is dropped after 10 if no action is taken
- Generally define minimum requirements
- Verify compliance: must come with a way of measuring whether the standard is met or not
- Something different from government – a standard has to be measurable and assessable
- Hierarchy of standards: code, standard, recommended practice, guide, policy
  - **Code** (e.g., building code) is the only legally enforceable standard. Only the government can make a standard a code.
  - **Standards** are sets of technical definitions and guidelines – how to documents

- **Recommended practices and guides** kept separate from standards but can evolve into higher-level document – recommended practice can evolve into a standard and a guide can evolve into a recommended practice.
- **Policies:** definition of roles and responsibilities, who has to do what, who is responsible for what. In the commercial sector, *policy* falls below *guide*.
- **Special Project report** is another document produced by standards groups. These are documents that support the others but don't require consensus.

In government terminology, a *standard* is a compromise – taking the least best (minimum acceptable); however, in the commercial sector, *standard* has a more specific meaning (e.g., safety standards, beta vs. vhs). Code, standard recommend practice, and guide are all consensus documents. Consensus means “substantial agreement (not an assigned percentage) – that shows a preponderance of the committee/group's support. Special project reports are not consensus. They are reviewed by committee and pertain to issues that should be included in one of the others.

All organizations agree to these definitions but they can be implemented differently by different organizations.

You create a standard by forming an ad hoc committee of one or more and survey the existing standards and societies to see if it already exists. If not, petition the board. They review and decide. The review board has all the power – there is no appeal.

**Question:** Is there a way for a person to complain that the committee isn't doing its job?

**Response:** Yes.

**Question:** Does everything evolve into a standard?

**Response:** No. Also note that the government [DoD] tends to order these terms differently and there often is no intention of having a guidance document evolve into a standard.

Historically, CFD guide published by AIAA, defines standards for CFD simulations. *Consensus* provides guidelines, but no recommendation regarding implementation. Journal of Fluid Dynamics, editorial guidelines that authors must address to publish their work. Examples: experimentalists must present error bounds together with their measurements. Simulations must demonstrate grid convergence (refer to Patrick Roache's grid convergence index).

- V&V programs generate evidence for validated simulations.
- Necessary for decision-making, that increasingly relies on simulations.
- V&V starts with the definition of requirements and the identification of risks.
- V&V balances the “cost of knowing versus risk of assuming”.

**Question:** Who's doing the V&V of the V&V? Where do we stop?

**Response:** Tools are being developed to evaluate software, judge the maturity of the software development of an organization, putting test procedures in place, define a methodology for the development of software. But this may not be the whole picture of V&V.

**Question:** Where do we get the experts to do validation?

**Response:** This is application specific. No single organization should do it within DoD because of the diversity of application and subject matters. Validation and accreditation are tied to the “intended use”, therefore, application dependent.

**Question:** What do we call a “model”? What about people who build hardware?

**Comment:** V&V should be an integrated part of the development process. Can it be defined as a checklist (“do A, B, C and be done)?

**Comment:** V&V should be done to assess the credibility of a simulation, to establish confidence that this model, hardware, piece of software meets requirements for the intended applications.

**Comment:** Credibility means the calculations I did were right for the right reasons.

**Comment:** Important to note for professional societies, “standard” requires a metric. Government should do this as well. We have a pretty detailed standard process – our metrics focus on whether this process is followed or not? metrics of what we need to see in the report. However, the level of what is done in there – we don’t have metrics. Government is looser – more high level. When you are talking a standard there is a specific of what people need to do to be successful – we tend to set out a structured approach for addressing the problem. On a global level, standard should be something all recognize as the way to go. Government does not provide a measurable way for people to know when they have met the standard.

**Comment:** Problem with nomenclatures. what is meant by the term standard is very soft. DoD use of term implies something different than it does in the commercial sector. People outside of DoD don’t necessarily relate to this – can be confusing.

Summary: Thoughts on future direction of V&V and standards:

- If this session had been held two years ago, we wouldn’t have had 3 people in the audience. There is greater emphasis on product liabilities, increased need for “certification” – product liabilities force greater need for testing, etc. if things aren’t standardized – can be very costly.
- The most important part is the validation. the code may come and go – the model generated by the code can have many different uses.
- If you don’t have a list of requirements, V&V is a waste of time.
- Validation metrics need to be meaningful comparisons.
- Need to collect a body of evidence because you can’t prove an absolute. Things are done by the preponderance of the evidence.
- Difference between IV&V and V&V – importance of the I – can be very expensive. Linda Rosenberg’s plenary presentation shows that IV&V does not replace V&V but is used for highly critical issues. IV&V can be extremely expensive, for example, development of a medical device, it may be 75% of the total cost. Independence lends weight to the evidence.

- Interesting quote – balancing the cost of knowing against the cost of assuming.
- Major concern that we have separated ourselves from M&S too well. We need to show we are a part of the whole process, that V&V is not a separate entity. We do V&V every day -- when shopping, buying a car, etc.
- Characterize verification in a broader sense – we were talking about code but this should also apply to requirements and design – this is where we have an opportunity to correct things before they get into the code. What do people think in policy areas? Certain guidelines and standards on verifying requirements and design can be very helpful.
- What we seem to be talking about is not V&V but about process. Perhaps we are as interested in process as in V&V – is important in understanding the context of the problem.

The second half of the session was devoted to presentations from DoD and the Services describing the relationships between DoD and Service VV&A policies and guidance. Each presentation followed the same outline

- M&S Management Structure
- VV&A Policy and Guidance
- VV&A Process
- VV&A Roles and Relationships
- VV&A Standards and Tools

**Presentation 2: “Department of Defense Modeling and Simulation VV&A Policy and Guidance,”** by Simone Youngblood (DMSO VV&A Technical Director)

One of DMSO’s primary functions is to develop policy. DoDI 5000.61 instructions policy defines the procedure for V&V, does not provide a specific explanation of what needs to be implemented. The VV&A Recommended Practices Guide (RPG) is a web-based series of documents developed by members of the DoD M&S community to provide guidance for conducting VV&A throughout DoD. Producing these documents generates a lot of discussion, helps for building consensus and helps everybody getting involved in the enforcement of these policies.

Underlying philosophy, principles and methodologies for VV&A in DoD was recommended in 1996 publication. Guidelines must account for diversity of audience, diversity of situations.

DMSO VV&A’s Web resource: <http://www.msiac.dmsomil/vva>

Depending on role (developer, user, V&V agent, manager, etc.) documents explain the roles and responsibilities in the V&V process. Does not, however, address implementation. VV&A processes differ, based on the type of model involved (legacy, new or federation). There are also differences based on type of application (training, analysis, acquisition) which are intended to be addresses in case studies. It has been extremely difficult to get people to come forward with lessons learned or case studies that can be used.

The term standard has some very specific connotations. It is the minimum information needed to communicate the policy. Recommended practice guides discuss how to implement policy and are not intended to evolve into standards. Also interested in lessons learned showing work-around solutions. There have been some very inventive ones (e.g., Joint Cruise Missile captured

necessary documentation using type of palm pilot and information categories and voice recorders (they found that removing the batteries destroys the memory so were acceptable in classified areas).

The DoD policies and guidance are intended to provide a foundation for each of the Services and Agencies to build upon. The next few presentations discuss the Service specific implementations of these.

**Comment:** Regarding the M&S Process diagram, if you don't understand the problem, you can't do VV&A.

**Comment:** Accreditation is not just a decision. It is a decision but is made as the result of an assessment.

**Comment:** There is still a heavy reliance on judgment. SMEs are extremely valuable throughout validation and assessment. [My] concern is that validation is always subjective versus objective.

**Response:** Efforts underway to become more objective by distancing observer from the process. One effort is focusing on ways for SMEs to evolve the referent first.

**Question:** When we come up with a policy, is it a standard?

**Response:** No, different services use the term *standard* differently. The term *standard* is a verification compliance statement. There are specific criteria involved in a standard. A recommended practice doesn't require such criteria.

**Presentation 3: “Department of the Navy Verification, Validation, and Accreditation Policy and Standards,”** Jennifer Park (NAVMSMO VV&A Lead)

**Key points:**

- Issues being explored: What is meant by “validation”? What's the reference to reality? How do you identify subject matter experts?
- The definitions used by the Navy are the same definitions throughout DoD.
- The focus of the Navy VV&A effort is on developing implementation guides. Policy can give you the recipe. The Navy is working on developing many specific implementations of the policy.

**Comment:** A conceptual model is validated not verified – not sure what this means. We still need a formal standard for conceptual model.

**Presentation 4: “Department of the Army Verification, Validation, and Accreditation Policy and Standards”** by Susan Solick

**Key points:**

- Army M&S policy designates AMSO as the focal point for VV&A.
- Army M&S policy recognizes class accreditation, accreditation of an M&S for use in a specific category of applications. The application sponsor would still be responsible for accrediting the M&S for the specific application. Class accreditation can be used on an

M&S being developed with no specific application in mind to focus the V&V effort and provide a foundation for future application accreditation efforts.

- Army M&S policy stipulates that VV&A is part of M&S configuration management.
- Army VV&A policy integrates user data V&V into the M&S VV&A process.
- The Army has a designated VV&A Representative (AMSO) responsible for policy, programs, education and a designated VV&A Standards Category Coordinator responsible for establishing standards and guidance
- The Army M&S Standards program focuses on establishing standards and identifying best practices, tools, and techniques in eighteen “areas of interest” in Army M&S, one of which is VV&A.
- Army M&S standards development is a formal, consensus-based process.
- The focus of the Army VV&A standards and guidance effort has been to tie into ongoing efforts to ensure consistency with other elements of DoD and compliance with Army policies such as, DoD VV&A RPG, IEEE 1278.4 (DIS VV&A), Data V&V, Levels of V&V.

**Presentation 5: “Department of the Air Force Verification, Validation, and Accreditation Policy and Standards” by Sam Johnson (USAF M&S Management)**

Identify needs, identify capabilities, propose standards.

**Presentation 6: “An Air Force Example” by Lynda Liptak (AFMC/DR, OAS)**

**T7 Session Participants (23)**

First Name	Lasr Name	Organization
Juan	Arvelo	Johns Hopkins University APL
Edward	Blankenship	Marine Corps
Alan	Breitler	
Grayden	Figart	Gray Hawk Systems, Inc
Cathryn	Francis	NNC Ltd
Francois	Hemez	Los Alamos National Laboratory
Sam	Johnson	Air Force M&S Policy Division
Jerry	Kniphfer	Virtual Targets Center
David D.	Leitch	JNIC (TRW Systems)
Lynda	Liptak	USAF OAS
Raymond	Lowman	Computer Sciences Corporation
William	Mccray	NAIC/TAND
Robert	McEniry	Office of Aerospace Studies
John	McVee	QinetIQ Marine Structure Dept
Megan	Meyer	Johns Hopkins University/APL
Linda	Olson	Acton Burnell, Inc.
Jennifer	Park	OPNAV/SSC San Diego
Philip	Parry	BAE SYSTEMS
Leonard	Schwer	Schwer Engineering & Consulting Servic
Susan	Solick	TRADOC Analysis Center
Marcia	Stutzman	Northrop Grumman
Andreas	Tolk	VMASC
Simone	Youngblood	DMSO