

Acceptability Criteria

Introduction

Although the literature on modeling and simulation (M&S) acceptability criteria is not sparse, as a recent paper noted, the literature has many limitations: “illuminating guidelines on the criteria themselves are rare.”¹ Also, it has been observed that considering the importance of acceptability criteria to M&S and M&S assessment, resources for acceptability criteria description “are miniscule compared to other related topics such as simulation requirements.”² This Special Topic was developed to help increase the understanding of how acceptability criteria should be identified and defined (described) and of the role that acceptability criteria play in verification, validation, and accreditation (VV&A). Such increased understanding of acceptability criteria may help to bring appropriate attention and resources to bear upon the development and use of acceptability criteria.

This Special Topic begins with a review of definitions for acceptability criteria and related terms. Both the official Department of Defense (DoD) definitions as well as terminology and definitions found in the M&S literature are considered so that connotations likely to be attached to acceptability criteria are recognized. This is followed by a discussion of the development of acceptability criteria: how they are identified, how they should be defined, and suggestions about formats used to describe them. The next section addresses assessment of acceptability criteria (i.e., judgment of the quality of the acceptability criteria) and determination of their satisfaction (i.e., assessment of whether the M&S verification and validation (V&V) evidence demonstrate that the M&S product satisfies the acceptability criteria). A section on using risk to tailor VV&A follows.

What Are Acceptability Criteria?

The official DoD definitions for acceptability criteria and related terms follow, included in the discussion are other connotations for these terms that the reader may encounter.

DoD Definitions

Official DoD definitions for accreditation, accreditation authority, and acceptability (or accreditation) criteria are found in the DoD policy for M&S VV&A, DoDI 5000.61³, and are presented below. Definitions in the M&S VV&A policy and guidance documents of the military Services are identical to or compatible with

these definitions. The DoD VV&A documentation standard for M&S, MIL-STD-3022⁴, also uses these definitions.

- **Accreditation** – The official certification that a model or simulation and its associated data are acceptable for use for a specific purpose
- **Accreditation Authority** – The organization or individual responsible for approving the use of models, simulations, and their associated data for a particular application
- **Accreditation Agent** – The organization designated by the M&S Application Sponsor (user) to conduct an accreditation assessment for an M&S application
- **Accreditation/Acceptability Criteria:** A set of standards that a particular model, simulation, or federation must meet to be accredited for a specific purpose

Acceptability criteria and accreditation criteria are synonymous terms. This paper uses acceptability criteria, but if one encounters the phrase “accreditation criteria” it should be understood to mean the same as acceptability criteria. For example, the 2009 version of the DoDI 5000.61 refers to “accreditation criteria” while MIL-STD-3022 refers to “acceptability criteria.” The March 2010 version of the DoD M&S Glossary defines the terms “acceptability criteria” and “accreditation criteria” as being interchangeable.⁵

Non-DoD standards and guidance closely related to DoD interests, such as the Institute for Electrical and Electronics Engineers Std 1516.4-2007 for the VV&A overlay to the High Level Architecture (HLA) Federation Development and Execution Process (FEDEP)⁶ or the U.S. Coast Guard Commandant Instruction 5200.40,⁷ use “acceptability criteria” in ways that are comparable to or the same as the DoD term of “acceptability criteria.” The North Atlantic Treaty Organization (NATO) M&S VV&A activities^{8,9} also use the DoD connotation for acceptability criteria when applied to M&S VV&A.

Characteristics of Acceptability Criteria

In the DoD, acceptability criteria are used to evaluate an M&S application for a specific intended use. That intended use should be agreed upon by those involved in the application domain where accreditation is desired.

Characteristics of acceptability criteria derived from the definitions presented earlier are:

- Acceptability criteria are always related to M&S accreditation assessment for a particular intended use. Note: there may be more than one intended use for which the M&S is being accredited.

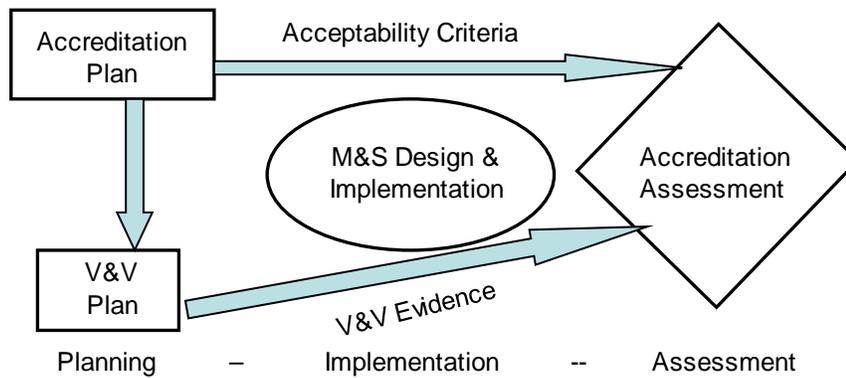
- Acceptability criteria should be approved by the Accreditation Authority. The criteria may be identified, drafted, tentatively defined, etc., by others (usually the Accreditation Agent), but until they are approved by the Accreditation Authority (normally in an Accreditation Plan), the criteria are not the official measuring stick against which the M&S is compared in the accreditation assessment.
- Acceptability criteria must be measurable such that an objective determination can be made as to whether or not the M&S satisfies the criteria.

Additional characteristics of acceptability criteria focus on how the criteria are developed:

- Acceptability criteria are directly related to the M&S requirements and provide the threshold against which a requirement will be assessed.
- Acceptability criteria are developed in conjunction with and are contained within the Accreditation Plan.
- Good acceptability criteria are clear (i.e., explicit, unambiguous); consistent (i.e., a requirement is not incompatible with and does not contradict any other requirement); and comprehensive (i.e., they address everything that the Accreditation Authority requires to make an accreditation decision).

Development of Acceptability Criteria

The relationship between V&V and Accreditation impacts the definition and application of acceptability criteria. While the acronym VV&A implies an ordering of processes with accreditation at the end, the accreditation process is implemented at both the front and back ends of V&V implementation. At the front end, the accreditation process is used to identifying the breadth and depth of V&V evidence that needs to be produced to support the accreditation assessment as well as the criteria used to assess against. At the back end, the accreditation assessment is conducted to determine if the M&S adequately meets the defined acceptability. Therefore, the acceptability criteria, defined in the early stages of a VV&A effort and documented as part of the Accreditation Plan, impact the implementation of V&V and the assessment of the V&V products. This relationship is illustrated below.



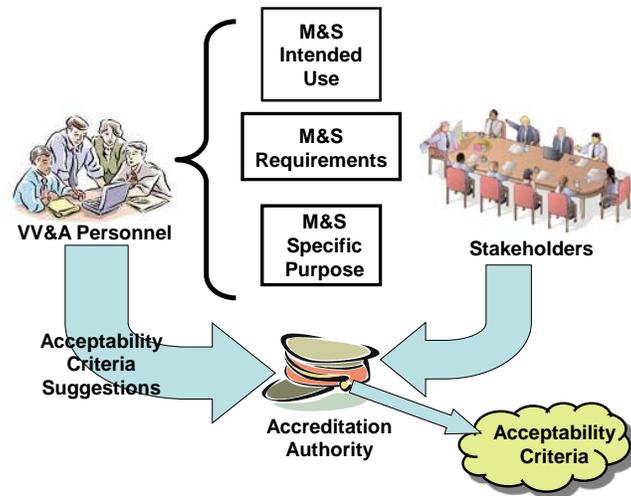
Logical and Temporal Relationships

When defining and documenting acceptability criteria, several key issues arise. Specifically, how are acceptability criteria derived? Who articulates the criteria? What format should be used to describe the criteria? This following section addresses these questions and others related to acceptability criteria development.

Deriving Acceptability Criteria

Methods for defining and deriving acceptability criteria vary widely as does the party responsible for deriving the criteria. This role may be undertaken by the Accreditation Agent, the Accreditation Authority, a User, a stakeholder, members of the VV&A team, or a combination of any or all these parties. The most common situation has the Accreditation Agent, working together with the Accreditation Authority, to draft the criteria. Regardless of who drafts the criteria, the Accreditation Authority has final approval authority.

Once the 'who' has been identified for the development of the acceptability criteria, the 'how' can begin to be addressed. The illustration below shows how the intended use for the simulation, the requirements for a new M&S development or M&S modifications, and the capabilities of a legacy simulation are all leveraged in the development of the acceptability criteria.



Identification of Acceptability Criteria

An important point to remember when deriving acceptability criteria is that the criteria are intrinsically linked to the M&S requirements. This relationship is discussed in more detail in the [Advanced Topics>Special Topics>Requirements](#). As indicated by the definition, acceptability criteria provide the standard against which the representational capability, as defined in the simulation requirements, is assessed. The acceptability criteria are also framed by the intended use, which dictates the required representational fidelity.

As noted by Harmon and Youngblood,² the acceptability criteria “need to be acceptable to all of the players in a simulation development, modification or application.” This can be facilitated by involving all application domain SME’s and the stakeholders in deriving the candidate acceptability criteria. The Accreditation Authority should consider the perspectives of all concerned when reviewing the candidate criteria. However, the Accreditation Authority has the final responsibility for setting the criteria and must be comfortable with them, since he or she is responsible for the assessment and recommendation.

Unfortunately, requirements and intended use statements may not have been documented or are inadequate. When requirements are not available, the VV&A team has to draft tentative acceptability criteria on the basis of the available information and interactions with the User, Accreditation Authority, and other stakeholders.

VV&A personnel must first determine the intended use and then scour available information about the M&S to support it. For a legacy simulation, documented requirements may no longer exist or at least may not be readily available. Descriptions of M&S capabilities may be available in an analyst manual for the simulation, or one may have to thoroughly review the code because available documentation explains only how to run the M&S and does not provide detailed

information about M&S capabilities. Obviously, existing V&V reports and assessments of the simulation are very helpful in the process, as are any requirements documents.

On the basis of the M&S information obtained, VV&A personnel create a set of tentative acceptability criteria that support the intended use for which accreditation is desired. During the process of developing this list of tentative acceptability criteria, VV&A personnel should interact with the Accreditation Authority, User, and other stakeholders to ensure that critical points from their perspectives are addressed by the tentative list of acceptability criteria. The tentative list of acceptability criteria should be presented to the Accreditation Authority for consideration. When approved, the tentative acceptability criteria become the acceptability criteria for the intended use.

The following statements should guide the development of the acceptability criteria.

- **Acceptability criteria are always related to an accreditation assessment of the simulation for a specified intended use.** Within the DoD definition, acceptability criteria are always within the context of M&S accreditation for an intended use.
- **Acceptability criteria within DoD should always be approved by the Accreditation Authority (or by the M&S User in the absence of an Accreditation Authority).** Criteria developed by VV&A personnel or others may provide useful insights about M&S capabilities, but without approval by the Accreditation Authority, such criteria should not be the basis for the accreditation assessment. That role is restricted to acceptability criteria that have been approved by the Accreditation Authority.
- **As far as possible, acceptability criteria should be derived from M&S requirements.** Typically, acceptability criteria will be directly related to M&S requirements when the M&S effort has good requirements. Additional elaboration may be required for some of the factors to ensure that the criteria are viable (i.e., they can be assessed within available VV&A resources) and their satisfaction can be determined unambiguously. Acceptability criteria not derived from good M&S requirements should be derived from the intended use for the simulation.

Parsimony in Acceptability Criteria Identification

The **law of parsimony** is “the logical principle that no more causes or forces should be assumed than are necessary to account for the facts.”¹⁰ Because resource limitations may prevent performing all possible V&V activities, parsimony should be applied to the identification of acceptability criteria. This will lead to the most efficient set of acceptability criteria necessary plan and

implement successful VV&A. Even with parsimonious acceptability criteria, it is likely that the acceptability criteria may need to be adjusted to balance the desired credibility in the accreditation assessment with available VV&A resource demands.

Acceptability criteria parsimony is achieved through consideration of a number of factors. First, acceptability criteria can be grouped so that decision-makers can address related criteria. An early writer on simulation acceptability assessment identified six major areas that need to be assessed in determining simulation acceptability¹¹:

- Simulation results
- Data
- Model
- Experimentation specification (statement of intended use, program)
- The code and M&S implementation
- Assessment methodology/technique

Often representational acceptability criteria are grouped topically. For example, all criteria related to sensor performance might be grouped together. It is important to remember, as noted by Harmon and Youngblood,² that those acceptability criteria may address more than representational characteristics. For example, processing speed may be an important criterion for M&S supporting real-time Live Virtual Constructive (LVC) exercises.

Second, since an acceptability criterion may be a measure of merit reflecting multiple measures of effectiveness that depend upon various measures of performance and that may be computed from lower-level measures, a single acceptability criterion can represent simultaneous acceptable performance of a number of M&S capabilities. For example, if in a missile or air defense/anti-air warfare simulation an acceptability criterion is a threshold on the probability of the attacking weapon (missile or bomb) reaching the target defended, satisfaction of that criterion would require successful performance by threat detection, tracking, and assessment processes as well as defensive weapon reliability, performance, and effectiveness in engaging the threat. In this situation, the VV&A personnel involved in the accreditation assessment would likely examine many factors, but the Accreditation Authority would only have to deal with the one acceptability criterion *if the criterion is satisfied*. If the criterion is not satisfied, the Accreditation Authority would want to know where the deficiency lay (root cause analysis) so that an appropriate decision could be made, such as no accreditation, a limited accreditation, or a decision to fix the problem and reassess the simulation.

Acceptability Criteria Definition and Description

MIL-STD-3022, the standard for VV&A documentation, includes acceptability criteria in each of the documents for which it provides templates, but MIL-STD-3022 does not specify how acceptability criteria should be described. That will vary with the type of simulation and its application. The specific guidance from MIL-STD-3022 relative to documentation of acceptability criteria is quoted below.

M&S REQUIREMENTS AND ACCEPTABILITY CRITERIA. This section describes the M&S requirements defined for the intended use, the derived acceptability criteria that should be met to satisfy the requirements, the quantitative and qualitative metrics used to measure their success, and the order of their priority. The relationships among the requirements, acceptability criteria, and metrics/measures can be shown either in the text or in a table (an example of which is shown below).⁴

Example Requirements Relationship Table

#	M&S Requirement	Acceptability Criteria	Metrics/Measures
1		1.1	1.1
		1.2	1.2
		1.n	1.n
2		2.1	2.1
		2.n	2.n
n		n.n	n.n

[From MIL-STD-3022, 2008]

The table above can be misleading. It implies a one-to-one relationship between an acceptability criterion and a metric or measure. Often a particular metric or measure will contribute to determination or satisfaction of more than one criterion, and a single criterion may involve many metrics and measures in determination of M&S capability to satisfy that criterion. The table (or description) of the relationship of requirements, acceptability criteria, and metrics/measures needs to be able to reflect the full set of potential relationships (which can include one-to-one, many-to-one, one-to-many, and many-to-many relationships among requirements, acceptability criteria, and metrics/measures).

MIL-STD-3022 also shows how acceptability criteria fit in the requirements traceability matrix, as illustrated in the table below. It is helpful if the numeration of acceptability criteria is related to the requirement (or requirements) to which the acceptability criterion is related. For example, if the acceptability criterion were related to requirements 3.2.6 and 14.3.1, the numeration identifying the acceptability criteria might be 3.2.6-14.3.1. Obviously this enumeration schema is neatest when the acceptability criterion is related to a single requirement. This schema also facilitates grouping acceptability criteria if the requirements are organized logically and topically.

M&S Requirements Traceability

	Accreditation Plan		V&V Plan	V&V Report	Accreditation Report
#	M&S Requirement	Acceptability Criterion	Planned V&V Task / Activity	V&V Task Analysis	Accreditation Assessment
1		1.1	1.1.1	1.1.1	1.1.1
			1.1.2	1.1.2	1.1.2
		1.2	1.2	1.2	1.2
		1.n	1.n	1.n	1.n
2		2.1	2.1	2.1	2.1
		2.n	2.n	2.n	2.n
n		n.n	n.n	n.n	n.n

[Table C-II from MIL-STD-3022, 2008]

Within this structure, the acceptability criteria should be organized so that they are grouped into logical topics and information about the relationship of each criterion to requirements and other criteria is specified.

The statement of each criterion should:

- Be clear, i.e., the criterion is neither ambiguous nor vague
- Be explicit about the conditions and circumstances associated with the criterion, so that satisfying or failing to satisfy the criterion will be clearly demonstrated
- Be consistent with the other criteria
- Provide links to background information such as pertinent M&S requirements, priority of related M&S capabilities, etc.

An “onion” structure for description of the acceptability criteria may be useful: the top layer consists of categories of criteria with overall measures of merit of that category; next follow detailed descriptions of criteria at various levels within a category; and finally background material (such as linkage to related M&S requirements, assessment processes, etc.) is included. Whether the onion structure is implemented by separate documents for the different layers or by a single document with flexibility in the level of information shown is not addressed here, because such decisions will depend upon the circumstances of each situation.

Acceptability Criteria Format

At present there are no formal standards for acceptability criteria description. None of the current formal M&S VV&A guidance from DoD, the military services and Defense agencies, or professional societies addresses in detail how acceptability criteria should be described or presented. Formal guidance

addresses acceptability criteria relationships to requirements, as do the few general statements of MIL-STD-3022 that were presented in the previous subsection, and may have general words about categorizing the criteria as essential or merely desired for the levels of performance, functionality, fidelity, or credibility indicated by the criteria. Therefore, format suggestions presented below should be viewed as suggestions that may be helpful, not as official guidance.

The volume of information that may be needed for acceptability criteria impacts formats that are acceptable. Tabular formats shown by the table extracts from MIL-STD-3022 in the previous subsections are appropriate for short names of acceptability criteria, but probably would be awkward for full descriptions of the acceptability criteria. Electronic tables, with linkage from a criterion name or numerical identifier to the full information of the criterion can make a tabular approach like the one shown in the tables of the previous section possible.

Regardless of the format, the five information items of an acceptability criterion typically are:

- 1) Identifier that relates to the requirements with which the criterion is associated. The identifier can also relate to V&V activities, tasks, and techniques, but such links would be developed as part of the V&V plan and are not part of the acceptability criterion itself. However, the acceptability criterion identifier may relate to items in the accreditation plan as suggested by the M&S Requirements Traceability Matrix of MIL-STD-3022.
- 2) Short name or description for the acceptability criterion.
- 3) Metrics or measures that will be used in evaluation for the criterion to determine if the M&S satisfies the criterion or not. This can be a qualitative assessment, as in a SME judgment that M&S results compare acceptably with the designated referent. This can be a single threshold value, such as the simulation must be capable of performing particular calculations within a set amount of time, or the results must not differ from the referent by more than a set amount. This can be a statistical measure of some kind. The essential aspect is that metrics or measures must be stated explicitly and contain enough description and explanation to prevent potential ambiguity.
- 4) Specified conditions of the evaluation. Normally this includes specification of the referent to be used (particular data, comparison with theory or results from a specified simulation running a specified problem, the sampling to be employed over the potential surface of referent information, etc.). It may also include the scenario(s) that are to be considered in which the criterion parameters and values are to be evaluated.
- 5) The method(s) to be used in performing the evaluation. If SMEs are to be used, their qualifications and processes may be specified. If particular statistical assessments are to be performed, the statistical techniques may be

specified. If particular V&V techniques are needed for assessment of the criterion, that should be stated.

Examples of Acceptability Criteria

The M&S product assumed in these examples of acceptability criteria is a new M&S development. Its purposes are to support analysis of the sensor suite for a semi-autonomous system at three phases of its life cycle: concept development and assessment, preliminary design, and operational use. The semi-autonomous system is intended to provide a capability for thorough inspection of the underside of cars, vans, trucks, trailers, and other such vehicles for detection of explosive devices attached to the vehicle while personnel are at a distance. It is expected to be capable of employment on normal roads, at check points, and in special facilities.

In a situation such as this, with the role of the simulation and its associated acceptability criteria extending from early in the M&S life cycle into its operational use, it is likely that the simulation will be modified during that time span and some of the acceptability criteria may also change. The examples of criteria discussed are assumed to remain the same throughout the M&S life cycle so that it can be shown how intermediate referent information (i.e., SME judgments) is specified before final referent information is available.

Two Illustrative Requirements

The examples of acceptability criteria defined and assessed in this Special Topic will relate to one or both of these requirements. The requirement alphanumeric identifiers are arbitrary and have no significance beyond allowing the acceptability criteria identifier to indicate its requirement relationship.

- **Requirement CP17:** The simulation shall be able to represent a coherent three-dimensional picture of the vehicle's underside from Semi-Autonomous System sensors and communicate the picture to the controllers, who are at a safe distance. The coherent picture shall be developed on the basis of information from all of the sensors to create an integrated picture of the underside of the vehicle.
- **Requirement P3:** The simulation shall be able to represent power usage for mobility, sensors, communication, and computation. The simulation shall be able to analyze future power expenditures for several scenarios (e.g., numbers and kinds of vehicles to be inspected) in order to support timing and planning for recharging of the batteries.

Illustrative Acceptability Criteria

Several illustrative acceptability criteria are defined below. Comments are presented about the acceptability criteria to explain some of the reasons the

criteria are as they are. Comments are given in a bulleted list below each example criterion.

Acceptability Criterion CP17.1 – Integrated Three Dimensional Picture

Development: The simulation must be able to produce a coherent three dimensional picture that deviates no more than 3 percent in any dimension from the actual three dimensional picture produced by the Semi-Autonomous System when using inputs of conditions from the test for which the referent picture was created. Satisfaction of the 97 percent accuracy requirements will be determined by comparing the simulation picture with the actual picture at intersection points on a grid of *x* evenly spaced lines across the picture horizontally and *y* evenly spaced lines across the grid vertically. Coherence of the picture will be determined by existence of only one line between any two points on the picture with no discontinuities in any line. Satisfaction of this criterion will be established for the set of vehicle types listed in scenarios A1-A4.

- This criterion has two parts. The first part is the capability to create a coherent three dimensional picture from multiple sensor inputs. It is likely that data from different sensors will produce different information about some parts of the vehicle's underside. Such differences may arise from differences in resolution and accuracy of the sensors, in their measuring different items (e.g., one sensor might ignore mud on a surface while another sensor measured it), from errors in the processes, or from other sources. The simulation, as well as the system itself, must be able to produce a coherent three dimensional picture in spite of such.
- The second part of the criterion concerns the accuracy of the three dimensional picture in reproducing what the system produces. The method for quantifying that accuracy is specified including the conditions (scenarios) for which it is to be measured. It is recognized that accuracy of the simulated three dimensional picture cannot be determined until the system exists and data from its testing are available. Accuracy judgments before that time will come from SME review.
- The criterion as stated provides for explicit and unambiguous determination of whether the criterion is satisfied. The criterion has an identifier that relates it to the requirement it addresses and the criterion has a short name. The criterion identifies metrics and measures to be used, along with conditions and methods for the evaluations.

Acceptability Criterion CP17.2 – Semi-Autonomous System

Communications: M&S representation of Semi-Autonomous System communication of the three dimensional coherent picture of a vehicle's underside to the controller must realistically replicate system capabilities. Simulated signal strength from the M&S representation of system communication with its controller will be measured every 10 degrees in azimuth for every vehicle type in scenarios A1-A4 for controllers at three specified distances with the controllers in the open

and in a bunker. Simulated signal strength must be within 3 decibels of measured signal strength from all pertinent data in related operational tests. Prior to the operational tests, assessment of M&S representation of signal strength will be based upon SME review, in which all pertinent test data or analysis will be used by the SME and included in explanation of the rationale for the SME assessment. Representation of the time required for computation of a three dimensional coherent picture and its communication to the controller must be within 25 percent of times experienced in related system tests. Prior to the system tests, SME judgment will be used to assess this aspect of M&S capability.

- The criterion specifies exactly what is to be measured and specifies referent information that is to be used for comparison. For this kind of criterion (one with multiple measurements), it is possible that most of the measurements would be acceptable, but some would not. As the criterion is stated, that would not satisfy the criterion. If it were desirable for such a situation to be acceptable, then the criterion would need to read “Simulated signal strength *on the average* must be within 3 decibels” instead of what it currently says (it currently omits *on the average*).
- The criterion has a short name and an identifier that relates it to the requirement it addresses. The requirement identifies metrics and measures to be used and the conditions and methods of the evaluation. In this, as well as in the previous example, the methods to be used address times in the SAS life cycle when test data exist and times before then. Sometimes additional detail is given about SME methods to be used.

Acceptability Criterion P3.1 – Semi-Autonomous System Sensor Power Analysis: The simulated representation of power usage for each sensor must not differ from development test results for the individual sensors by more than 10 percent when simulated conditions reflect test conditions. Sensor power usage will be analyzed for every valid development test. Before test data are available, satisfaction of this criterion will be based upon SME review and estimation.

- This criterion is simple and straightforward. As stated, any failure of simulation results to match test data prevents satisfaction of the criterion. If simulation results matched in every case but one test condition (such as a test involving both high temperature and high humidity) for one sensor, it might be desirable to consider that acceptable. In which case, the criterion should be restated by adding *in no more than x situations for an individual sensor* at the end of the criterion.
- The above comment and the similar one for the previous criterion show the importance of being clear about what is essential for the simulation to be acceptable.
- The criterion has a short name and an identifier that relates it to the requirement it addresses. The criterion addresses metrics and measures,

conditions of the evaluations are specified, and the method of evaluation is implied (simple accuracy computation).

Acceptability Criterion P3.2 – Scenario Power Consumption: The simulated representation of power consumption must agree to within 20 percent of test results for each scenario in the set of scenarios tested in system operation tests. Prior to that time, evaluation of this criterion will be based upon SME review. SME review will include reference to all pertinent development tests and analysis with inclusion of such SME assessment reports.

- This criterion is explicit and its determination will be unambiguous. Note that the criterion appears consistent with the other criterion related to the P3 requirement. The accuracy of representation of individual sensor power usage is higher than the accuracy required for power usage by the system. How uncertainties should be combined is a complex subject and varies with the dependencies among components of the item. Typically the uncertainty percentage for the composite is smaller than the uncertainty percentage of the individual component.
- There is also a kind of consistency between this criterion and criteria related to the CP17 requirement. Referent information for them is to come from the operational tests when that stage of development is reached. Using a common source for referent information is an aspect of criteria consistency.
- The criterion has a short name and an identifier that relates it to the requirement it addresses. The criterion identifies metrics and measures, conditions of the evaluation, and methods to be employed in the evaluation.

Normally, acceptability criteria are approved as an entire set in an accreditation plan and not for each individual criterion unless the criteria have been approved piecemeal instead of being treated as a whole. The approval indication should identify the authority approving the criteria and the date of the approval, and should indicate where the approval is documented.

Acceptability Criteria Assessment

Quality of the Acceptability Criteria

Assessment of acceptability criteria involves checking to confirm the following:

- Criterion approval by the Accreditation Authority for each criterion. This assessment should be based upon available documentation, not just based upon verbal statements.

- Viability of acceptability criteria, both as a collection of criteria (they can be assessed and their satisfaction determined) and individually (no criterion requires information or application of techniques not likely to be available during the assessment period). This assessment normally will be a judgment by the Accreditation Agent, V&V Agent, and/or designated SME(s).
- Proper articulation of the acceptability criteria. They are clear and unambiguous to everyone involved in the accreditation assessment, their satisfaction can be clearly determined, they are consistent, they are organized logically, they are explicitly related to intended uses and M&S requirements, and each criterion is described by the five information items noted in the previous section. This assessment normally will be a judgment by the Accreditation Agent, V&V Agent, and/or designated SME(s).
- Acceptability criteria scope and rigor are compatible with credibility levels desired by the Accreditation Authority. If the risk-based VV&A methodology is being employed, more specific questions will be addressed, such as: Will acceptability criteria scope and rigor produce acceptable levels of M&S use risk for the intended use? This determines if all essential functional and non-functional aspects are addressed by the criteria with enough assessment rigor to support the level of M&S use risk required by the User. This assessment requires application of risk-based VV&A methodology during accreditation and V&V planning to ensure that parameters used in calculating M&S use risk and relative to the quality of V&V information seem compatible with an accreditation assessment that is based upon the acceptability criteria. Any V&V activities or tasks in the V&V plan that do not seem to be related to any of the acceptability criteria should be noted and evaluated to see if they impact M&S use risk or if they merely consume V&V resources that otherwise might be applied to acceptability criteria. This assessment should be performed under the supervision of the Accreditation Agent and the V&V Agent.

Assessment of the acceptability criteria requires a good bit of information beyond the criteria themselves. The assessment requires some level of information about the people involved in the accreditation assessment and the accreditation in order to judge the clarity of the descriptions of the criteria for them. For example, if statistical comparisons are expected, a specific statistical test (or tests) to be used should be named. Knowledge of the personnel would determine whether simply naming the statistical test is adequate or if some description of the test is also required for clarity of the criterion.

Assessment of quantitative criteria is straightforward because the criteria should be explicit with regard to the methods for their computation and the acceptable limits of the evaluation. Assessment of qualitative criteria is prone to ambiguity and uncertainty. For clarity and to remove ambiguity, the qualitative criteria may

prescribe how the qualitative assessments are to be done. For example, a criterion might indicate the report format and kinds of content expected with qualitative assessments. Use of a SME form established for M&S assessments that requires both the SME judgment and conclusion *and* the rationale upon which the conclusion is based might be noted in the criterion as a form to be used in the assessment.

If the evaluation of acceptability criteria reveals any deficiencies, they should be corrected before the criteria are presented to the Accreditation Authority for approval.

Acceptability Criteria Satisfaction

Some criteria can be assessed simply on the basis of functionality. For example, a criterion that requires the simulation to consider specified factors, to employ a specific methodology or technique, or to possess a capability to merge diverse elements into a single entity can be assessed simply on the basis of the simulation's functionality.

Other criteria need referent information in their assessment. The best referent is high-quality test data for well-understood conditions with complete documentation of uncertainties associated with the data and its circumstances. The next best referent are data from tests that may not be of high quality, or for which test circumstances are not completely known, and the items in the test are similar but not identical to those modeled in the simulation. Often uncertainties for such data are not completely known. The least reliable referent sources are theory, SME estimation (expert judgment), results from similar simulations, etc.

Methods for comparing simulation results with referent information vary. The most reliable and credible are formal methods that provide mathematically provable conclusions. The next most reliable and credible methods employ standard statistical techniques that quantify uncertainties and statistical likelihoods for the conclusions reached. The least reliable and least credible methods are the equivalent of "eyeballing" correspondence between two curves on a chart. Nearly all SME assessments fall in this category.

In addition to assessment of individual acceptability criteria, it is important to step back and ensure that criteria are being assessed in an appropriate context and that all major concerns of the User and Accreditation Authority have been addressed by the acceptability criteria. This is a necessary sanity check for any accreditation assessment. It also is important for acceptability criteria since their primary function is to support accreditation assessment.

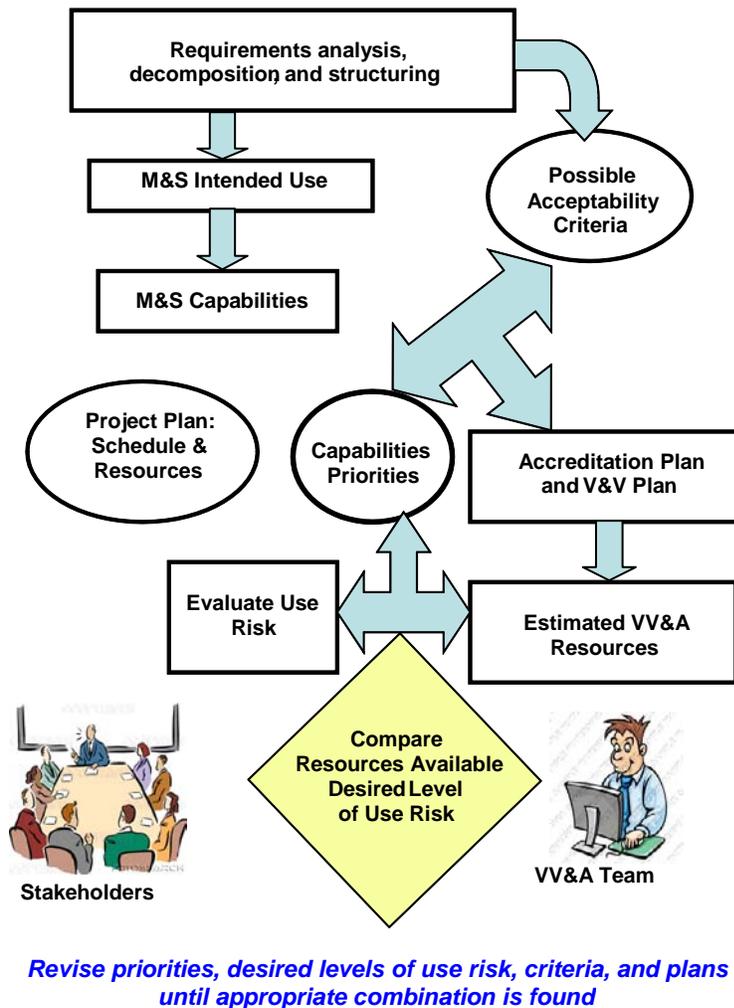
As the accreditation assessment is performed in accordance with the accreditation plan, a scoreboard of acceptability criteria should be maintained. It should indicate the status of each criterion. Status entries include the following possible entries:

- Assessed per accreditation plan (date, report) and fully satisfied.
- Assessed per accreditation plan (date, report) and not fully satisfied. Description of consequences:
 - Simulation disqualified for intended use
 - Corrective action underway; to be reassessed (estimated date)
 - Accreditation Authority decided to modify the criterion (eliminate it, change it so it is satisfied, or declare the performance such that it only limits M&S use in specified ways)
 - Becomes limitation on Intended Use
 - Other
- To be assessed per plan (no problems expected or noted)
- To be assessed per plan at new date (with potential problems, such as referent unavailable at present)
- Assessment to be modified from that of the accreditation plan (describe how and why; also give new assessment date). Data that are close to passing may be reviewed and limits/criterion extended if this does not adversely affect the results for the intended use.
- Removed from assessment for specified reason (such as schedule or resource limitations)

Scoreboard information should be available to the Accreditation Agent and V&V Agent and to others at their discretion. The scoreboard is a principal tool for assessing the acceptability criteria and will contribute valuable information for the accreditation report.

VV&A Planning and Risk-based VV&A

It is important that the scope of the V&V and accreditation efforts are consistent and that the V&V evidence produced is focused in the areas that the accreditation assessment will ultimately be performed. The reality of any V&V and accreditation effort is that the availability of resources will impact both the planning and implementation phases of the process. Resources constraints, whether cost, schedule, personnel, or equipment related will have an impact on the scope of an effort.



The VV&A Planning Process

During the VV&A planning process, depicted in the figure above, the acceptability criteria, anticipated VV&A resources, priorities of M&S capabilities relative to the intended use, and acceptable levels of M&S use risk are addressed so that a suitable combination is determined. Sometimes adjustments have to be made to capabilities, priorities, or acceptable levels of use risk, and sometimes additional resources (or schedule) are necessary.

The process illustrated above begins with M&S requirements. They are analyzed and then decomposed into lower-level requirements that support the higher-level requirements initially stated. Part of this process includes supplementing the requirements with derived requirements, as described in [Advanced Topics>Special Topics>Requirements](#). The full set of requirements is structured so that they are grouped topically, and the numbering system for them is structured so that related requirements have related numbers. From the

analysis, decomposition, and structuring of the requirements come two things: identification of M&S capabilities pertinent to the simulation's intended use, and possible acceptability criteria. As the V&V and accreditation plans are developed, the resources required can be estimated. If the resources for a full VV&A effort exceed what may be available, options should be defined. This may involve limiting the acceptability criteria and/or requirements that will be addressed in the VV&A effort, applying additional resources, changing tolerable levels of use risk, or some combination of these things.

For small projects, the VV&A planning process proceeds in an ad hoc fashion. For larger or more visible projects, a formal risk-based VV&A methodology has been developed to facilitate tailoring of the VV&A plans. The risk-based VV&A methodology is a process for interacting with M&S stakeholders to determine prioritization of M&S capabilities relative to simulation's intended use. The Accreditation Agent and V&V Agent will develop an accreditation plan and a V&V plan to support the accreditation assessment. These plans will be developed with knowledge of acceptable levels of M&S use risk to the User. From the plans, required VV&A resources can be estimated. The risk-based VV&A methodology calculates M&S use risk during the planning phase on the basis of priorities assigned to the capabilities, tolerable levels of M&S use risk, and VV&A resources expected to be applied.^{12,13,14,15,16}

The risk-based VV&A Methodology consists of four key steps with multiple activities¹⁵:

“1.0 Determine Needed Capabilities. Describe required capabilities (in terms of objects, dependencies, and the required accuracy of those representations).

- 1.1 Obtain Requirements
- 1.2 Identify Capabilities
- 1.3 Group Requirements

“2.0 Characterize Risk Areas. Estimate tolerable levels of use risk in terms of effects on the user decision and capability importance, and then assign those estimates to different parts of the required capabilities to reflect the nature of the intended use.

- 2.1 Specify Effects of M&S Use on User Decision
- 2.2 Determine Capability Importance
- 2.3 Determine Capability Evidence Maturity Level

“3.0 Plan V&V. Develop the V&V Plan by choosing V&V tasks, techniques, and coverage according to the assigned capability importance levels.

- 3.1 Choose Activities and Tasks
- 3.2 Choose Technique

- 3.3 Estimate Schedule, Cost, and Resource Requirements
- 3.4 Estimate Confidence

“4.0 Collect and Report V&V Evidence. Successively refine importance assignments and V&V tailoring decisions as more information is gained and the V&V needs evolve.

- 4.1 Characterize M&S Capabilities
- 4.2 Characterize M&S Limitations
- 4.3 Communicate M&S Use Risk Evidence to Users”¹⁵

Risk-based VV&A provides a rational process based upon sound mathematical principles for establishing relationships among M&S use risk (related to mistaken belief about the correctness of simulation results); the seriousness of mistaken use of simulation results; the quality of V&V evidence as it impacts accuracy of the assessment of simulation capabilities; and the quality of V&V evidence that can be produced by application of V&V activities, tasks, and techniques at the level of simulation capabilities. The risk-based VV&A process permits rational decisions about selection of V&V activities, tasks, and techniques (i.e., tailoring them for assessment) so that the impact of V&V resource usage on the quality of V&V evidence produced and its consequences for correctness of assessment can be quantified. As such, the V&V and accreditation plans can be designed to optimize VV&A investments.

Conclusion

This Special Topic has presented the DoD definition for M&S acceptability criteria and provided context about the concept of acceptability criteria so that other connotations for the term are known and their relationship to the DoD definition understood. It also provides information about how to identify acceptability criteria, how to define them, and how to assess them. A five-element format is suggested for description of acceptability criteria (identifier, short name, metrics/measures, conditions, and methodology) with illustrative examples of acceptability criteria. This Special Topic also shows how acceptability criteria play in the risk-based VV&A methodology.

Acceptability criteria should always be the basis for accreditation assessment in DoD M&S accreditation decisions. As more emphasis is placed upon modeling and simulation in DoD activities, credibility of M&S results gains increasing importance. Proper acceptability criteria provide a sound basis for credibility of results from accredited M&S products.

Acceptability criteria need to be defined early in the M&S life cycle for new M&S developments or modifications. Since logically efficient and effective V&V is designed to support prospective accreditation of the simulation for its intended

use *and* to ensure that the M&S satisfies its requirements, acceptability criteria need to be defined in the early stages of accreditation planning and before V&V planning to support it are final. Hence, acceptability criteria will be established early in the M&S project.

For application of a legacy simulation, acceptability criteria need to be defined early in the project for the M&S intended use that is to undergo an accreditation assessment. This is necessary for efficient use of V&V resources to generate the information needed to support the accreditation assessment in a timely way.

If acceptability criteria are not defined early enough, the accreditation assessment may be delayed while requisite information is generated to support it.

References

1. Campbell, James D., "Defining Acceptability Criteria: the Key to the VV&A Process," *Spring 2010 Simulation Interoperability Workshop (SIW)*, paper 10S-SIW-032, 12-16 April 2010.
2. Harmon, Scott Y., and Simone M. Youngblood, "Lessons Learned about Developing Acceptable Acceptability Criteria," *Proceedings of the 2008 European Simulation Interoperability Workshop*, paper 08E-SIW-056, 16-19 June 2008.
3. Department of Defense (DoD) Instruction 5000.61, *DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)*, 9 December 2009.
4. Department of Defense (DoD) Standard Practice MIL-STD-3022, *Documentation of Verification, Validation, and Accreditation (VV&A) for Models and Simulations*, 28 January 2008.
5. Department of Defense (DoD), *DoD Modeling and Simulation (M&S) Glossary*, March 2010, available at http://www.msco.mil/files/Draft_MS_Glossary_March_B_version.pdf.
6. Institute of Electrical and Electronics Engineers (IEEE) Std 1516.4-2007, *Recommended Practice for Verification, Validation, and Accreditation of a Federation – An Overlay to the High Level Architecture Federation Development and Execution Process*, 20 December 2007.
7. U.S. Department of Homeland Security United States Coast Guard Commandant Instruction 5200.40, *Verification, Validation, and Accreditation (VV&A) of Models and Simulation (M&S)*, 22 December 2006.

8. Masys, A. J., M. L. van Emmerik, and P. Bouc, "Verification, Validation and Accreditation (VV&A) – Leveraging International Initiatives," in *Transforming Training and Experimentation through Modelling and Simulation* (pp. 10-1–10-10). Meeting Proceedings RTO-MP-MSG-45, Paper 10. Neuilly-sur-Seine, France: RTO, 2006, available from <http://www.rto.nato.int/abstracts.asp>.
9. Masys, A. J., M. Roza, C. Giannoulis, and J. Jacquart, *Verification, Validation, and Accreditation (VV&A): The GM-VV Contribution and Roadmap*, RTO-MP-MSG-060, undated.
10. *Oxford English Dictionary*, 2nd edition (1989) with additions series (1993–1997).
11. Oren, T. I., "Concepts and Criteria to Assess Acceptability of Simulation Studies: A Frame of Reference," *Communications of the ACM*, 24 (4), April 1981, pp180-189.
12. Johns Hopkins University Applied Physics Laboratory, *The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Literature Survey and Methods Bibliography*, NSAD-R-2009-207, October 2009.
13. Johns Hopkins University Applied Physics Laboratory, *The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Assessment of Applicable Methods*, NSAD-R-2010-001, January 2010.
14. Johns Hopkins University Applied Physics Laboratory, *An Approach for Realizing a Risk-Based VV&A (risk-based VV&A) Methodology*, NSAD-R-2010-020, March 2010.
15. Johns Hopkins University Applied Physics Laboratory, *An Approach for Realizing a Risk-Based VV&A (risk-based VV&A) Methodology Comments Summary*, NSAD-R-2010-020, September 2010.
16. Johns Hopkins University Applied Physics Laboratory, *Risk-Based Methodology for Verification, Validation, and Accreditation, The M&S Use Risk Methodology*, NSAD-R-2011-011, April 2011.

Acronyms

DoD	Department of Defense
DoDI	Department of Defense Instruction
FEDEP	Federation Development and Execution Process
HLA	High Level Architecture
IEEE	Institute of Electrical and Electronics Engineers
M&S	Modeling and Simulation
MIL-STD	Military Standard

RPG	Recommended Practices Guide
SME	Subject Matter Expert
V&V	Verification and Validation
VV&A	Verification, Validation and Accreditation

The appearance of hyperlinks does not constitute endorsement by the DoD, M&S CO, the administrators of this website, or the information, products or services contained therein. The DoD does not exercise any editorial control over the information found at these locations. Such links are provided consistent with the stated purpose of this website.