NSAD-R-2011-011

TECHNICAL REPORT

Risk Based Methodology for Verification, Validation, and Accreditation (VV&A)

M&S Use Risk Methodology (MURM)

April 2011

FOR:
Office of the Assistant Secretary of Defense,
Research & Engineering
Systems Engineering/Systems Analysis
1901 N. Beauregard St. Suite 500
Alexandria, VA 22311-1705

BY:
The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Road
Laurel, MD 20723-6099

This page intentionally left blank.

The Johns Hopkins University Applied Physics Laboratory Team

Title	Name	Organization
Team Lead	Simone M. Youngblood	JHU/APL
SME (VV&A)	Marcy Stutzman	Northrop Grumman
SME (VV&A)	Dale K. Pace	Independent Consultant
SME (Risk)	Peter P. Pandolfini, PhD	JHU/APL

This page intentionally left blank.

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The Deputy Assistant Secretary of Defense for Systems Engineering [DASD(SE)] identified his priorities for Fiscal Year 2011 (for further information see http://www.acq.osd.mil/se/). At the top of the list is:

- Support the current fight; manage risk with discipline.

The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Project was tasked with defining an integrated methodology for planning, tailoring, and conducting VV&A of model(s), simulation(s), and the associated data (hereinafter referred to collectively as "M&S") that seeks to provide insightful information that contributes to minimizing the risk that users incur when applying M&S to their intended uses. The purpose was to establish a disciplined approach to identify the risk involved in using M&S. Managers must be aware of the risks they are incurring when relying on M&S (i.e., abstract representations of the real thing) to provide information.

The objective of this risk-based methodology is to optimize VV&A resource use while minimizing the risks of using an M&S. The VV&A resources include the information available to support the gathering and analysis of verification and validation (V&V) evidence, the personnel required to execute VV&A activities and tasks, and the time available within which to perform the VV&A activities.

The information resources include the development products available for V&V, the best available relevant knowledge about the things being modeled or simulated, the description of the user needs for the M&S, and any supporting information needed, including germane regulations, standards, and guidance. The M&S Use Risks arise from many factors, including uncertainties in the representations of the M&S, uncertainties in the V&V evidence itself, and the consequences that result from using an M&S for an intended use. The uncertainties, if left uncharacterized, could lead the M&S User to either believe M&S Results are correct when they are not or judge M&S Results as incorrect when, in fact, they are correct. Either of these two errors could lead to adverse consequences for the M&S User or those whom they represent.

This methodology aims to avoid adverse consequences by informing M&S Users of limitations that might lead to either type of M&S use error. However, VV&A efforts generally must operate within constrained resources. The reality of resource constraints requires tailoring VV&A processes to optimize resource use within those constraints, to accommodate the alignment of VV&A with the M&S life cycle, and to address the specific priorities of an intended use. Tailoring a V&V effort involves choosing which activities and tasks to perform to collect the V&V evidence, what techniques to use to execute those activities and tasks, and what parts of the M&S behavior space to explore (i.e., sampling). All of these choices directly affect V&V effort costs. All tailoring decisions must balance VV&A resources required against the benefits provided in the reduction of M&S Use Risks.

An initial approach for defining a risk-based methodology was delivered in March 2010. This initial approach was vetted across multiple communities and with the risk professionals at the Software Engineering Institute, and was matured into a disciplined and rigorous methodology with a strong mathematical underpinning. The resulting product is the M&S Use Risk Methodology (MURM).

The MURM prioritizes and tailors M&S assessment based on M&S Use Risk analysis and documents the findings. The methodology provides a mechanism for identifying in advance specific potential problems with applying M&S capabilities relative to the intended use and communicating those problems to stakeholders. Implementing MURM identifies the key risks associated with the use of M&S Results to inform the intended use. It aligns planning and implementation of verification and validation activities with the most critical M&S Use Risks. Additionally, it builds stakeholder consensus around what is to be verified and validated (and how much is needed) and what is not to be verified and validated (and why).

Program Managers, M&S Users, Accreditation Authorities — whatever the name given to the role with the responsibility for making the decision to use M&S — can use a disciplined approach to identify the specific risks associated with using an M&S for an intended use. When communicated to stakeholders, M&S Use Risk information contributes to managing and mitigating the programmatic cost, schedule, and performance risks associated with implementing VV&A; thus, addressing DASD(SE)'s top priority of providing a disciplined approach for managing risk.

1. INTRODUCTION

This document is the fifth in a series of technical reports produced by The Johns Hopkins University Applied Physics Laboratory (JHU/APL) in performance of the Department of Defense (DoD) Modeling and Simulation Steering Committee (M&S SC) sponsored High Level Task V-AQ-2. The purpose of the High Level Task was to foster cost-effective verification, validation, and accreditation (VV&A) by developing a risk-based methodology for model(s), simulation(s), and associated data (hereinafter referred to collectively as "M&S"). The first four technical reports available from the DoD Modeling and Simulation Coordination Office (M&S CO) [Reference (a)] are:

- The Risk-Based Methodology for Verification, Validation, and Accreditation Literature Survey and Methods Bibliography, JHU/APL NSAD-R-2009-207 [Reference (b)].
  - This report documented the initial assessment of the state of the art for risk methodologies focused on surveying resources and publications to identify existing applicable methods.
- The Risk-Based Methodology for Verification, Validation, and Accreditation Assessment of Applicable Methods, JHU/APL NSAD-R-2010-001 [Reference (c)].
  - This report described the results of the assessment of methods documented in JHU/APL NSAD-R-2009-207 [Reference (b)] that were found applicable.
- An Approach for Realizing a Risk-Based VV&A Methodology, JHU/APL NSAD-R-2010-020 [Reference (d)].
  - This report described an initial approach to defining an integrated methodology for planning, tailoring, and conducting VV&A to minimize the risk that users incur when applying a model or simulation for its intended use.
- An Approach for Realizing a Risk-Based VV&A Methodology – Review and Comments Summary, JHU/APL NSAD-R-2011-018 [Reference (e)]

  - This report summarized the comments and feedback received from reviewers of the initial methodology presented in NSAD-R-2010-020 [Reference (d)].

1.1 BACKGROUND

Modeling and simulation is a more critical technology for DoD than ever before. The volume of M&S uses has never been greater nor have the ramifications of decisions informed by applying M&S Results been more far-reaching. By definition, M&S are abstractions of the real world and, using M&S to inform decisions, carries with it risk, specifically, M&S Use Risk[1] that is defined as follows:

> *The probability that inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker.*

Therefore, the need to efficiently and effectively analyze and assess that specific type of risk by applying a systematic M&S Use Risk Methodology (MURM) is critical.

It should be noted that M&S Use Risk is not programmatic risk (i.e., cost, schedule, and performance risk), but it can be impacted by all three.

Moreover, it must be stressed that in today's world of uncertainty, decision-makers need quick and easy access to M&S Use Risk information for better decision-making. Applying MURM will enable decision-makers to:

- Utilize a visually understandable scorecard of disparate data to make implementation and use decisions based on timely and accurate VV&A status information
- Identify the risks associated with a given M&S in mathematically cogent ways and develop insights about how to select the most appropriate verification and validation (V&V) methods
- Make defensible V&V tailoring decisions that enable effective and efficient use of V&V resources

The appropriate level of effort associated with VV&A is directly related to the impact of results from M&S capabilities on a decision, as well as the criticality of that decision. Determining the importance of the decision and the dependence on

[1] The derivation for this expression is found in Appendix F.

M&S Results in making the decision together set the stage for implementing MURM.

V&V are systems and software engineering process areas focused on assessing an M&S throughout its lifecycle. V&V are implemented to provide the evidence necessary to gain knowledge about M&S assumptions, capabilities, and limitations in relationship to the acceptability criteria. V&V leverages not only systems engineering and software engineering, but also information science, the cognitive and behavioral sciences, and other associated disciplines.

Accreditation is the process that identifies the acceptability criteria and applies the evidentiary knowledge gained from implementing V&V to determine the impacts to the intended use. This partitioning lets the Accreditation Implementation focus on both the intended use and the requirements for <u>using</u> the M&S and lets the V&V Implementation focus on collecting the evidence to be used in the accreditation assessment.

In a resource-constrained environment, the cost of verifying and validating M&S capabilities can be high. Cost of V&V can be increased by the absence of information from development products, historical VV&A documentation, or appropriate referent data. Therefore, it is important that V&V investments be weighed against the risk of making a bad decision because of unreliable M&S Results or because of uncertainty in the V&V assessment of M&S capabilities. Decision-makers responsible for VV&A investments should consider the following: the likelihood that capabilities of an M&S are flawed; one or more of the M&S software elements, hardware components, or data are flawed; or the M&S capabilities are misunderstood by the M&S User.

However, not all the software, hardware, or data comprising an M&S represent equal risk. A disciplined application of MURM can help identify the most important aspects of the M&S to be examined during V&V and provide insights into the relative risks of the various elements of an M&S. Because MURM is integrated with VV&A planning, it can help optimize VV&A investments.

MURM will help VV&A practitioners identify the important aspects of the M&S that when verified and validated will provide the evidence needed to mitigate the risk involved with using M&S Results to support decision-making. Decision-makers should use MURM over the M&S lifecycle to identify and track M&S Use Risk, make tailoring decisions to focus VV&A resource allocations, and make decisions applying the M&S Results.

This document is not a primer on the generic VV&A processes, rather it assumes the reader has basic knowledge of why VV&A is needed, what each process is (and is not), how each process relates to the other, what happens in each process, VV&A roles and responsibilities, and the particular problems associated with domains. Readers who want information about VV&A should refer to the DoD VV&A Recommended Practices Guide (RPG) [Reference (f)], which is freely available online at http://vva.msco.mil/. MURM overlays and complements the basic VV&A processes.

Information produced from the application of MURM is captured in the four core VV&A documents described in Military Standard (MIL-STD)-3022 [Reference (g)]:

- The Accreditation Plan focuses on defining the acceptability criteria to be used during the accreditation assessment; defining the methodology to conduct the accreditation assessment; defining the resources needed to perform the accreditation assessment; and identifying issues associated with performing the accreditation assessment.
- The V&V Plan focuses on defining the methodology for scoping the V&V effort to the application and the acceptability criteria; defining the V&V tasks to that will produce information to support the accreditation assessment; defining the resources needed to perform the V&V; and identifying issues associated with performing the V&V.
- The V&V Report focuses on documenting the results of the V&V tasks; documenting M&S assumptions, capabilities, limitations, risks, and impacts; identifying unresolved issues associated with V&V implementation; and documenting lessons learned during V&V.
- The Accreditation Report focuses on documenting the results of the accreditation assessment and documenting the recommendations in support of the accreditation decision.

Production of the four core VV&A documents is integrated in MURM.

1.2 DOCUMENT ORGANIZATION

This document builds upon the groundwork laid in the first four project deliverables. Chapter 1 provides an introduction and describes the contents found throughout the rest of the report. Chapter 2 provides the core information about MURM. Chapter 3 summarizes the depth and breadth of information presented in the report and makes recommendations for improving the methodology though future users' feedback.

The appendices contain a listing of references, descriptive information about the V&V Composite Model (VCM), the V&V Process Maturity Model (VPMM), the V&V Techniques Catalog, the substantive proof of the mathematical logic behind the disciplined methodology, and a listing of abbreviations and acronyms.

This page intentionally left blank.

2. OVERVIEW

MURM is not just a collection of tables and figures (as presented in other risk-related products). MURM applies that information and translates it into an underlying mathematical formula based on the definition for M&S Use Risk as provided in Section 1:

> *The probability that inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker.*

Therefore, M&S Use Risk is the probability that the inappropriate applications are true, that the unacceptable consequences are true, and that the unacceptable consequences are a result of the inappropriate applications. To put the definition into a mathematical form that can be used in a numerical calculation of M&S Use Risk, the definition was parsed into the following statement:

The probability that [(inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker) **AND** (that inappropriate application of M&S Results for the intended use occur) **AND** (unacceptable consequences to the decision-maker occur)].

In symbols an explanation of the mathematical notation can be found in Appendix G):

Causes = C ≡ inappropriate application of M&S Results for the intended use

Effects = E ≡ unacceptable consequences to the decision-maker

which results in the equation:

$$\text{M\&S Use Risk} = p[(C \wedge E) \wedge (C \Rightarrow E)]$$

where

$$p(\text{Causes}) = p(C_1 \cup C_2 \cup C_3)$$

with,

C_1 ≡ Lack of clarity of intended use leading to misuse (i.e., Clarity),

C_2 ≡ Adverse impact on decision if capability is not achieved (i.e., Importance), and

C_3 ≡ Incorrect recommendation to employ or not to employ M&S Results relative to that capability (i.e., Confidence).

And where

p(Effects) is the probability of the effects resulting from unacceptable consequences. p(Effects) is based on M&S Impact and M&S Reliance.

2.1 MURM CHARACTERISTICS

MURM has several characteristics that distinguish it from other processes that consider risk in M&S assessments. These characteristics are:

1) MURM is based upon a disciplined and rigorously coherent mathematical process so that its definitions of M&S Use Risk and other terms have explicit mathematical expressions, and those mathematical expressions are logically consistent and coherent. Methods used have been chosen both for mathematical coherence and to avoid unintended bias (that often creeps into processes) in weightings of factors. The rigorous translation of technical terms into mathematical expressions will facilitate automation of many aspects of the M&S assessment process. Presentation of MURM in this report goes to great length to ensure the reader has grasped the solid mathematical foundation upon which the methodology is based so that tools will find appropriate acceptance and use when they are developed.
2) MURM is described in general terms because it is a methodology that should be capable of application to any M&S no matter the category, type, domain, or application.
3) MURM is intended to be used by all stakeholders and practitioners associated with either accreditation or V&V efforts. It is expected that MURM will be adopted and used routinely with every M&S so there will be less risk of misuse of M&S capabilities in decision-making.
4) In addition to the mathematical rigor of its processes and concept definitions, MURM builds upon and reuses other previously published DoD-sponsored VV&A research and credible risk resources.
5) MURM has multiple purposes that include: (1) to provide a disciplined and mathematically cogent method for assessing risk from M&S use; (2) to facilitate effective and efficient use of V&V resources through the rationale for V&V tailoring; and (3) to provide a mechanism to effectively communicate M&S Use Risk to the M&S User. All of the above impacted the mathematics of the methodology and the purposes are accommodated by MURM.

2.2 METHODOLOGY DESCRIPTION

The description of methodology in this section will explain how MURM would be implemented as a part of the VV&A processes. Additionally, the underlying mathematical logic will ultimately enable automation of the methodology to support more efficient implementation and to streamline the production of information in easily communicated formats.

Figure 2-1 provides a graphic depiction of the MURM overlay to the VV&A processes. The descriptive overview that follows begins in the top left corner with the M&S Intended Use and M&S Requirements, continues down through Accreditation Planning, V&V Planning, V&V Implementation, and then up through Accreditation Implementation and the Accreditation Decision until it terminates in the upper right corner with the actual use of the M&S.

Throughout the explanation of the methodology, the related appendices and specific sections of appendices will be referenced. Additionally, the various parts of the MURM overlay to the VV&A processes will also reference the detailed discussions.

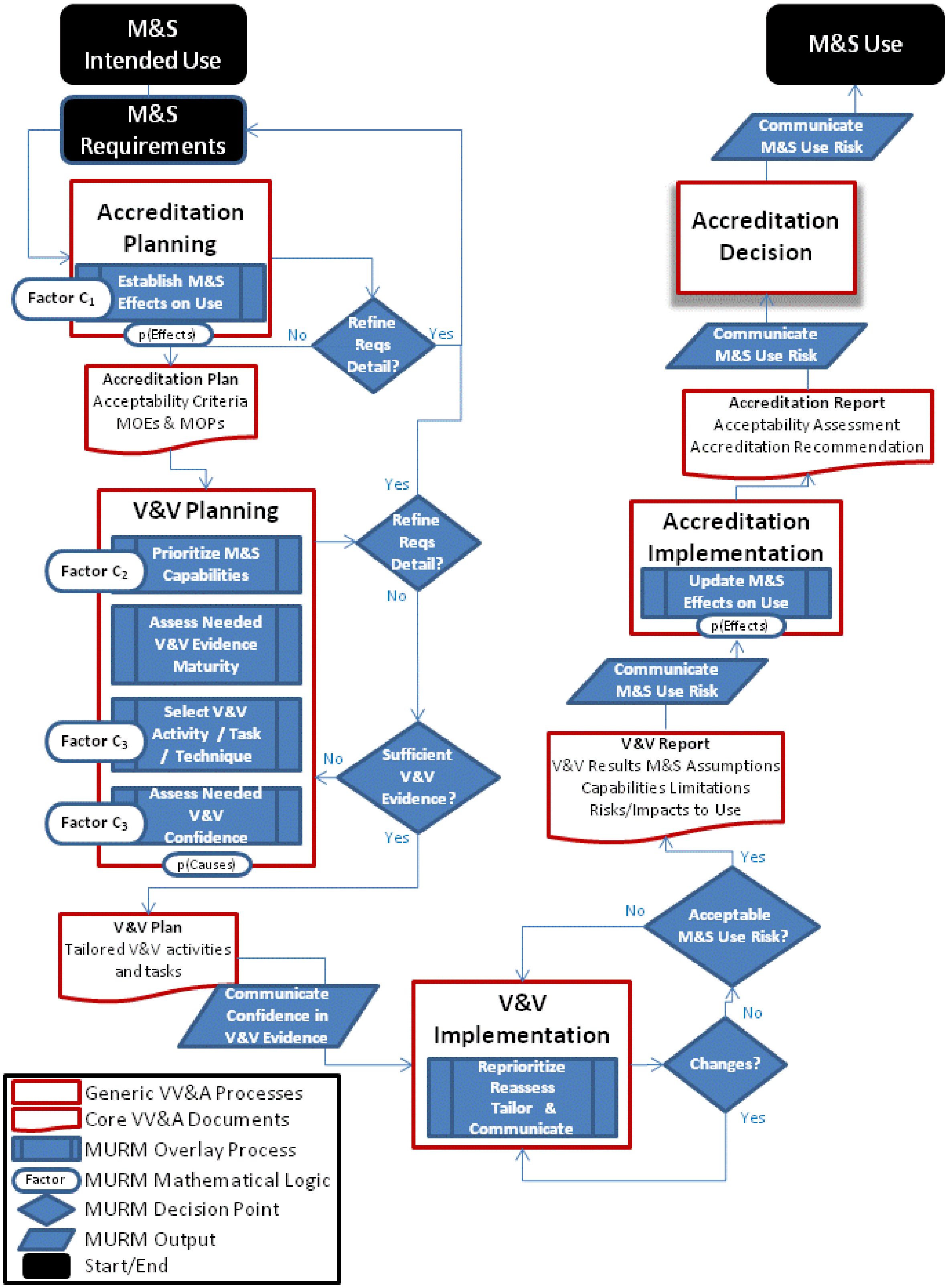


Figure 2-1: MURM Overlay to the VV&A Processes

2.2.1 M&S Intended Use and M&S Requirements

No matter the type of M&S to be used or the reason that it requires VV&A to be performed, the VV&A processes must begin with a statement of the M&S Intended Use and the specification of the M&S Requirements. Providing this information is the responsibility of the stakeholders – the people with the need to use M&S to support their program, mission, objectives, analysis, experiment, training, etc.

The assumption of a clear, complete, and concise M&S Intended Use statement and detailed M&S Requirements is not always realized. Rather, the clarity of the M&S Intended Use statement and the specification of the M&S Requirements are often not defined in sufficient detail and further refinement is necessary throughout the VV&A processes to ensure requirements are valid and testable.

Stakeholders must validate and approve all refinements of the M&S Intended Use statement and the M&S Requirements specification. The clarity of the M&S Intended Use and M&S Requirements directly affects the calculation of M&S Use Risk [see discussion about Factor C_1 (Clarity)].

Accreditation Planning as it relates to MURM is described in the next subsection.

2.2.2 Accreditation Planning

The Accreditation Authority represents the M&S User in the VV&A processes. The M&S Intended Use and M&S Requirements together represent the information needed by the Accreditation Authority (often represented by an Accreditation Agent) to initiate Accreditation Planning and produce the Acceptability Criteria and the associated measures of effectiveness (MOEs) and measures of performance (MOPs).

The M&S Intended Use, M&S Requirements, Acceptability Criteria, and associated MOEs and MOPs are all documented in the Accreditation Plan and traceability is maintained throughout the subsequent documents (V&V Plan, V&V Report, and Accreditation Report) [see Reference (g)]. The detail included in the Acceptability Criteria and associated MOEs and MOPs directly impacts V&V Planning, V&V tailoring, and M&S Use Risk.

2.2.2.1 M&S Impact and M&S Reliance

At this point in the application of MURM, M&S Impact and M&S Reliance are identified and the p(Effects) component of the M&S Use Risk definition (see Appendix F4) is determined.

p(Effects) is the probability of the effects resulting from unacceptable consequences to the decision-maker.

M&S Impact and M&S Reliance should be considered when stakeholders are thinking about the impact of M&S on program objectives and beginning to make decisions about program resources and program risk.

- M&S Impact is an indication of how much information the M&S is providing relative to the decision space.
- M&S Reliance is an indication of the dependence on using M&S Results in making the decision.

The M&S Intended Use statement is the key information used to identify M&S Impact as shown in Table 2-1.

Table 2-1 and Table 2-2 provide notional examples of how a stakeholder (e.g., Program Manager, M&S User, or Accreditation Authority) initially determines the programmatic risk incurred by deciding to use M&S.

Table 2-2 shows the stakeholder's dependence on using the M&S Results. Whether a stakeholder depends solely on the M&S Results or has other sources of information determines M&S Reliance.

It is important to note the numbers in the left columns of Table 2-1 and Table 2-2 (as well as the numbers used in Figure 2-7) represent a numerical ordering rather than weighted values.

Table 2-1: M&S Impact

M&S Intended Use	M&S Impact
5	**Intended Use** addresses **multiple areas of high impact to the decision,** key experiment, study, or analysis; key program review or test event; key system performance analysis or requirements definition; primary test objective or test article design; critical operational issue; key technical or managerial decision; critical skills training; regulatory compliance, licensing, permitting, or law.
4	**Intended Use** addresses **a single area of high impact to the decision**, key experiment, study, or analysis; key program review or test event; key system performance analysis or requirements definition; primary test objective or test article design; critical operational issue; key technical or managerial decision; critical training; regulatory compliance, licensing, permitting, or law.
3	**Intended Use** addresses **multiple areas of medium and low impact to the decision**, other experiment, study, or analysis, other program review or test event; other system performance analysis or requirements definition; secondary test objective; other skills training; other technical or managerial decision.
2	**Intended Use** addresses **a single area of medium impact to the decision**, other experiment, study, or analysis, other program review or test event; other system performance analysis or requirements definition; secondary test objective; other skills training; other technical or managerial decision.
1	**Intended Use** addresses **a single area of low impact to the decision,** objective or analysis that is not a significant factor in the technical or managerial decision-making process.

Table 2-2: User Reliance on M&S in Decision-making

M&S Reliance	
4	**M&S** will be the ***only*** method employed to support the decision-making process.
3	**M&S** will be the ***primary*** method, employed with other non-M&S methods, to support the decision-making process.
2	**M&S** will be a ***secondary*** method, employed with other non-M&S methods, to support the decision-making process, and will provide significant data unavailable through other means.
1	**M&S** will be a ***supplemental*** method, employed with other non-M&S methods, to support the decision-making process, and will provide supplemental data already available through other means.

Figure 2-2 and Figure 2-3 show how the information in Table 2-1 and Table 2-2 is used to help a stakeholder decide not only to do VV&A, but also provides an indication of the rigor of the V&V evidence that is needed to support an accreditation decision. For a more detailed explanation of the mathematical logic associated with Figure 2-2, see Appendix F4.

	M&S Impact				
M&S Reliance	Single Low Area	Single Medium Area	Multiple Medium or Low Areas	Single High Area	Multiple High Areas
Only	3	4	4	5	5
Primary	2	3	4	4	5
Secondary	2	2	3	4	4
Supplemental	1	2	2	3	4

Figure 2-2: Example M&S Impact and Reliance Matrix (p(Effects))

Organizations use matrices as visualization tools. The number in each cell of Figure 2-2 reflects the combination of M&S Impact and M&S Reliance information. The use of colors in matrices provides additional visual cues that are very familiar to stakeholders. The colors in Figure 2-2 reflect in a relative manner the cost and risk trends depicted in Figure 2-3.

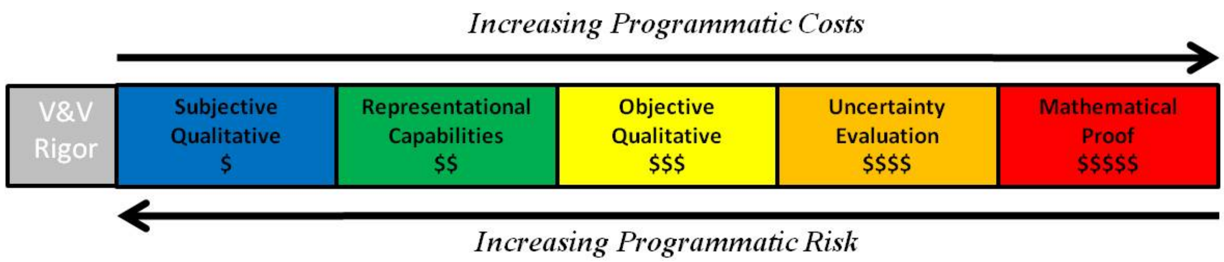


Figure 2-3: V&V Rigor

Figure 2-3 depicts V&V Rigor.[2] In the figure, the upper arrow shows the trends of the costs to the program to achieve more rigor, while the lower arrow emphasizes the decreasing M&S Use Risk to the program. The sequence of the activities, from left-to-right, is based on the suggested rigor of the V&V evidence needed to produce the specified level of evidence for comparison with the Validation Referent. The rigor of evidence is based on research to define maturity levels for organizations involved in implementing V&V processes (see Appendix C). It spans from the application of the most subjective techniques [e.g., simple face validation by a Subject Matter Expert (SME)] to the application of the most objective techniques (e.g., mathematical and statistical proofs).

A relative cost is assigned by the dollar signs ($) in each box to give a stakeholder an idea of the resources that might be needed relative to the programmatic risk being addressed.

As can be seen by the bottom arrow, Programmatic Risk decreases as the rigor of the V&V evidence increases, while the M&S Use Risk increases as the rigor decreases as shown by the top arrow.

[2] This section addresses the relative impact to a program when M&S, in general, is used. To decrease programmatic risk associated with using M&S, increased V&V rigor is needed. Appendix C addresses the V&V rigor needed to increase confidence in the accreditation recommendation to apply a specific M&S for an intended use, and thus decrease program risk when the M&S Results inform decision-making processes.

As stated previously, programmatic risk can impact the determination of M&S Use Risk. An important example is schedule (i.e., a plan of action and milestones). The schedule lays out the timeline and important milestones for accomplishing all Accreditation activities and tasks and can have a major impact on the scope of the VV&A efforts.

2.2.2.2 Clarity of Intended Use (Factor C_1)

An understanding of the M&S Intended Use, as well as a clear and concise determination of M&S Requirements and the associated Acceptability Criteria, is critical drivers for the implementation of effective and efficient VV&A. Additionally, a well-defined referent with minimal uncertainty is key to increasing confidence in validation results. Factor C_1 (Clarity) in the M&S Use Risk equation is computed based on the completeness of all of these elements, as shown in Figure 2-4.

The information used for determining Factor C_1 (Clarity) is provided by the stakeholders (i.e., Intended Use and M&S Requirements) and the Accreditation Authority (i.e., Acceptability Criteria, associated MOEs and MOPs, and referent).

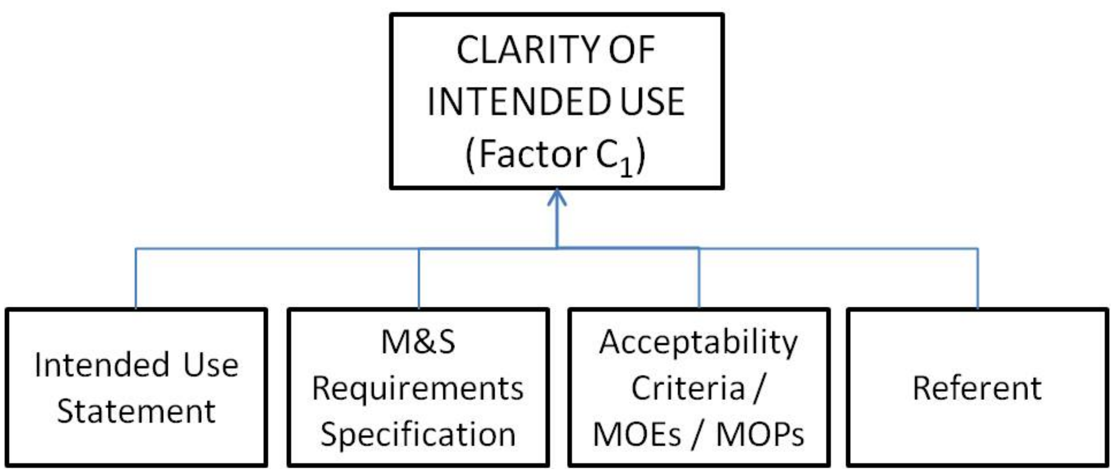


Figure 2-4: Clarity of Intended Use (MURM Factor C_1)

As expected, an imprecise or unclear definition of M&S Intended Use or M&S Requirements would negatively impact VV&A implementation and introduce additional uncertainty in the calculation of M&S Use Risk.

For discussions about determining Clarity of Intended Use (Factor C_1) see Appendix F, and for a more detailed example see Appendix F5. V&V Planning as it relates to MURM is described in the next subsections.

2.2.3 V&V Planning

MURM supports the development of a V&V Plan tailored to the available resources and schedule as defined by the Accreditation Plan. V&V contributes to reducing M&S Use Risk by reducing the uncertainties in the M&S User's knowledge of the M&S capabilities and limitations, thereby reducing the possibility of making use errors. However, complexity and resource limitations may make tailoring of the V&V effort necessary to meet realistic cost and schedule constraints.

Tailoring a V&V effort involves choosing which activities and tasks to perform to collect the V&V evidence, what techniques to use to execute those activities and tasks, and what parts of the M&S behavior space to explore (i.e., sampling). All of these choices directly affect V&V effort costs.

V&V Planning begins with determining needed capabilities within the M&S undergoing Accreditation. The M&S Intended Use, M&S Requirements, Acceptability Criteria, and associated MOEs and MOPs documented in the Accreditation Plan are continued in the V&V Plan [see Reference (g)]. V&V is planned to provide the evidence needed to address the acceptability criteria and within the schedule of planned actions and milestones established in the Accreditation Plan.

To characterize M&S Use Risk areas, tolerable levels of M&S Use Risk are estimated in terms of the effects on the M&S User's decision and M&S Capability Importance, and then those estimates are assigned to different parts of the required capabilities to reflect the nature of the intended use.

M&S Capability Importance defines the impact to the M&S Intended Use if a capability is or is not represented by the M&S. The strength and coverage (rigor) of the V&V evidence determines the uncertainty in the knowledge of M&S capabilities and limitations. Tasks and sampling choices determine the coverage of the V&V evidence. Maturity of the technique determines the strength of the V&V evidence. All of these choices affect the amount of effort required to obtain V&V evidence.

A correlation exists between rigor and cost, which translates to a correlation between risk and cost.

2.2.3.1 Tailoring

MURM employs tailoring to facilitate effective and efficient use of VV&A resources in determining M&S capabilities and their credibility.

MURM employs three levels of tailoring. The first level of tailoring occurs when determining priorities of the M&S capabilities within the scope of the overall V&V effort as defined by M&S intended use.

The second level of tailoring occurs in selection of V&V activities and tasks to be performed. Tailoring here is based upon which activities and tasks are viable because of information available or expected to be available. Activities, tasks, and combinations of tasks are ranked according to the quality of information they can produce (information quality is measured in accordance with the VPMM (see Appendix C).

The third level of tailoring occurs in selection of the V&V techniques to be employed. Detailed tailoring at this level addresses selection of the V&V technique that produces information with requisite quality in a manner that is consistent with available V&V resources.

Tailoring at these three levels is discussed briefly in the following subsections.

2.2.3.2 Tailoring Level 1: Identifying M&S Capability Priority and VV&A Scope

Often schedule and resources for M&S assessment (i.e., VV&A) are inadequate to permit addressing all requirements and achieving the rigor desired. Therefore, it is important to prioritize M&S capabilities so that V&V can be scoped to focus on the capabilities most important to the intended use.

Basic systems or software engineering principles are applied to M&S development or modification. Requirements to build or modify M&S are identified, transformed into a conceptual model, various specifications and design drawings, and then implemented into software and hardware components. V&V are the engineering practices implemented and documented to ensure requirements are traceable throughout intermediary development products and across the M&S lifecycle. The end result of the M&S development or modification processes is acceptance-testing leading to acceptance of the delivered product by the stakeholder. When an M&S is to be applied for a specific intended use and will require accreditation, additional actions are initiated. The requirements for the use of the M&S are identified and compared against the requirements to build or modify along with the information in the available development products. The output of those comparisons is gap information. Specifically, the V&V information gaps indicate additional V&V evidence that needs to be generated.

Whether it is for M&S development, modification, or use, VV&A resources generally will be limited. This in turn limits the scope and depth of the V&V effort. Focusing on the M&S capabilities that are most critical to the intended use through prioritization will provide the richest set of evidence to support either an acceptance or an accreditation decision. There are multiple ways to prioritize M&S capabilities (e.g., binning, structured relationships, etc.).

If an organization does not have an established methodology for prioritization, the following examples define some methodologies that can be employed.

Example methods for prioritizing M&S capabilities are shown by the two paths in Figure 2-6 and depend upon the quality and completeness of the information available in the Accreditation Plan. With well-formulated M&S Intended Use, M&S Requirements, Acceptability Criteria, and the associated MOEs and MOPs, the V&V effort can estimate consequences and determine M&S Capabilities Importance (Path B). With little or no accreditation information available, then M&S Capabilities Importance can be assigned based on consequences only (Path A).

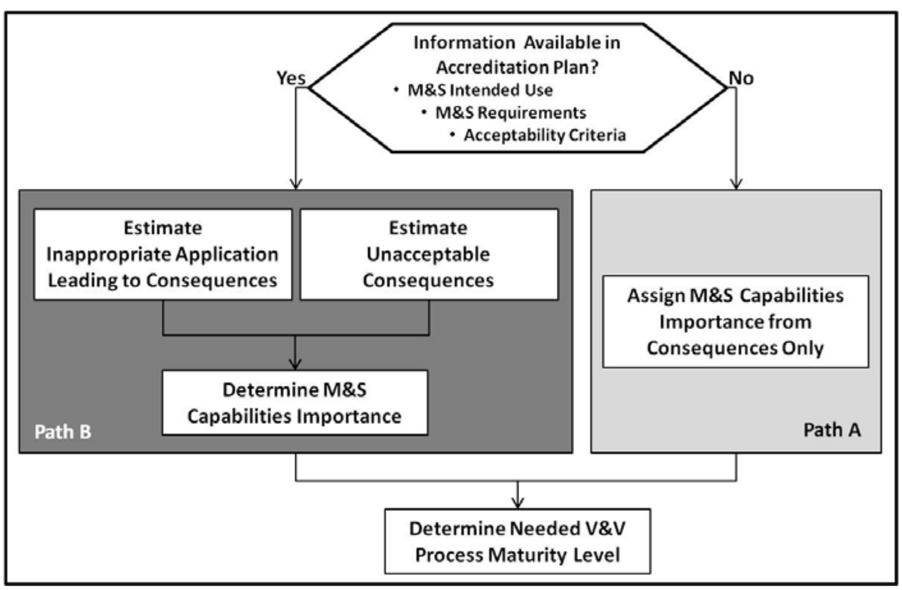


Figure 2-5: Paths to Describing M&S Capabilities

The information depicted in Figure 2-7, which represents Path B in Figure 2-6, includes representations of limitations, mitigation, and probability of unacceptable consequences, all of which are factors that are considered in computing Factor C_2 (Importance). For detailed information about the mathematical logic, see the discussion in Appendix F concerning the importance of the capability to the decision and in Appendix F3 concerning the analysis for the importance factor state table.

	A limitation will result in NEGLIGIBLE consequences	A limitation will result in MINOR consequences	A limitation will result in SERIOUS consequences	A limitation will result in GRAVE consequences
Mitigation Not Required	1			
Mitigation Complete	1	1	2	2
Mitigation Partial	1	2	2	3
Mitigation Impossible	2	2	3	3
	Probability of Unacceptable Consequences (low, medium, high ■■■)			

Figure 2-6: MURM Factor C_2 (Importance)

The essential elements or features of the M&S can be characterized into objects, attributes, behaviors, functions, algorithms, and metrics, etc. These elements may be documented in various development products making it easier to identify them. If they are not, then VV&A practitioners should break down the M&S into a structure that captures the essential elements and communicates the importance of each capability to the intended use. Involving the M&S Developer, if available, is essential to accomplishing it in an efficient manner.

Figure 2-8 provides an example of how to group M&S capabilities using object and dependency abstraction hierarchies based on what is important to the M&S User (i.e., the Intended Use). This grouping simplifies the communication of required capabilities to a picture. This simplification helps to focus the V&V attention on the relevant areas and plays a very important role in optimizing V&V resource use.

In this notional example, the intended use is to analyze the capabilities of the A-10 Thunderbolt jet aircraft to put laser-guided bombs on target during air-to-ground warfare operations. The M&S has many capabilities, some of which are not important for the specific intended use. The M&S structure is depicted both object and dependency hierarchies. The importance of the capabilities is identified by the numerals 1, 2 and 3, with a "3" indicating a high importance, a "2" a moderate importance, and a "1" a low importance.

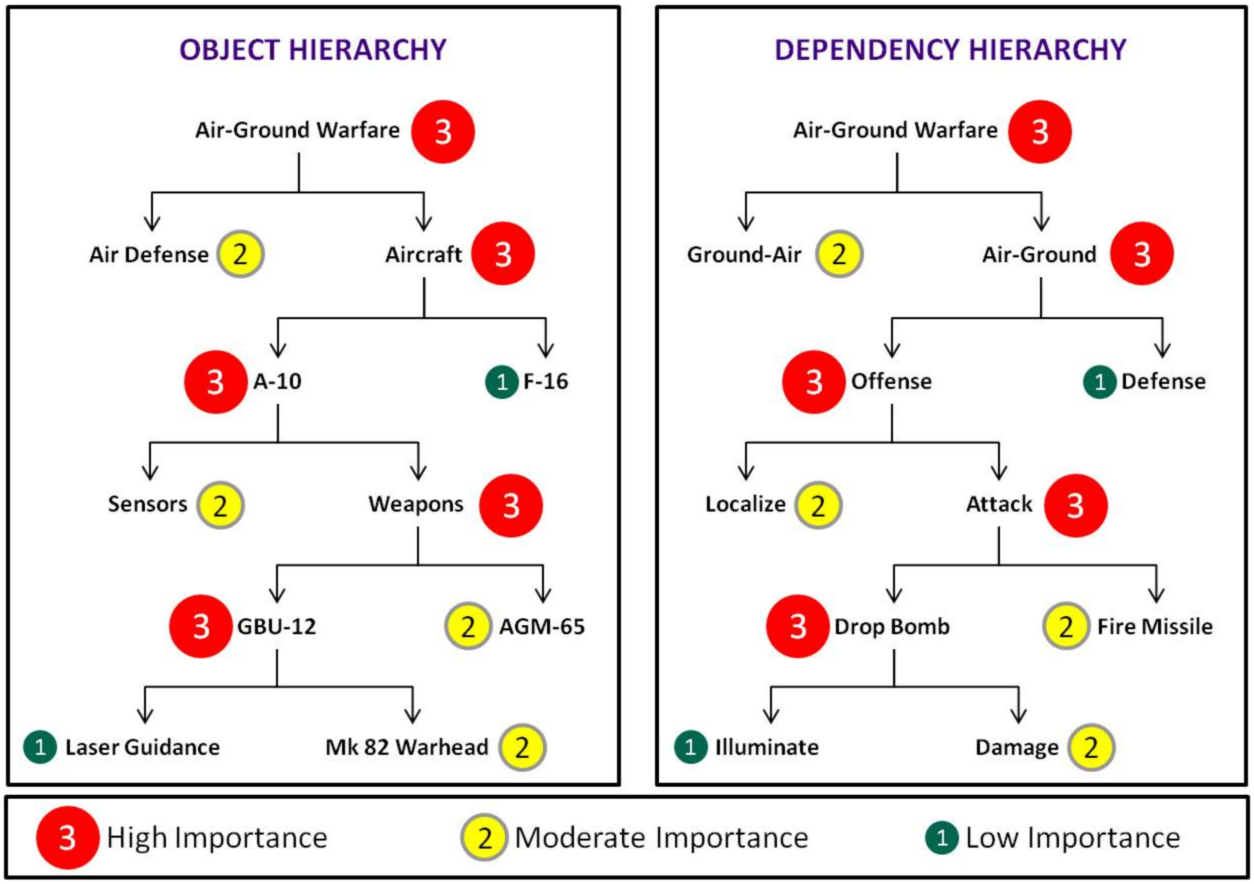


Figure 2-7: M&S Capability Importance

The higher the importance means there could be more risk involved in applying the M&S Results for the specified intended use if that capability was not represented in the M&S or was incorrectly represented. Once identified as high, medium or low, then the capabilities at each importance level must be prioritized. The result should be a prioritized listing of the most (down-to-the-least) important elements of the M&S. This is a method to communicate to the M&S User and other stakeholders where to focus the V&V evidence collection activities in proportion to the significance of those capabilities to the intended use.

Communication of the priorities is important because the M&S User and other stakeholders must make resourcing decisions based upon the priorities and available V&V resources. Once decisions are made, then VV&A planning can continue, estimates of the required resources and schedule for the highest priorities can be determined, and M&S Use Risk assigned.

Adjustments in priorities and risk will be made as other factors affect the schedule and resources; these are to be expected in any program, experiment, or study. The goal is to deal with the unexpected when it happens and to provide the M&S User and other stakeholders with a re-prioritized capabilities listing and the resources required to address the most important capabilities at that time.

Once an organization knows where they need to focus V&V resources, then they can decide how to tailor V&V activities to obtain the necessary V&V evidence.

At this point in MURM, Factor C_3 (Confidence) is computed based on tailoring of the V&V activities and tasks to be performed and the selected techniques discussed in the following subsections. See Appendix F for detailed discussions about computing Factor C_3 (Confidence).

2.2.3.3 Tailoring Level 2: Selecting V&V Activities and Tasks

Tailoring and ranking of viable V&V activities and tasks are based upon information availability or expected availability as described in Appendix B. In this level of tailoring, possible V&V activities and tasks are drawn from the VCM (see Appendix E). The resulting list of ranked viable V&V activities and tasks is updated as the VV&A effort progresses. If time and resource limitations do not permit performance of all viable V&V activities and tasks, it will be necessary to restrict tasks so that minimum acceptable quality information is produced for the higher priority capabilities, stopping when resources or time limitations are reached unless modifications are made to resources, schedule, or intended use.

2.2.3.4 Tailoring Level 3: Selecting V&V Techniques

In performing V&V activities and tasks identified, selection of the V&V technique is driven by the quality of information required for the task, resources and time required to perform the technique, and capabilities of personnel available to perform the V&V tasks. V&V techniques are identified in the V&V Techniques Catalog (see Appendix D).

The information as defined below is needed to make informed V&V technique selection:

- Possible V&V Techniques

  - The V&V Techniques Catalog (Section 2.2.3.5 and Appendix D) identifies a collection of V&V techniques that should be adequate to produce whatever assessment information quality is possible relative to a particular M&S capability. The Catalog indicates the quality of information produced by a technique and has some indication of resources and personnel expertise required to employ the technique. The Catalog is expected to be a living document, with new techniques added as they come into use (such as V&V techniques that may be developed for use with adaptive simulation processes).

- Referent Information
  - Availability and quality of referent information sets bounds on the quality and rigor of assessment information that can be produced.
- Personnel expertise and availability
  - Different V&V techniques require different levels of personnel expertise. Personnel availability is an important consideration, especially in regard to SMEs who may be needed for particular aspects of the assessment.
- M&S Resource and Schedule
  - The V&V Plan should contain information about VV&A resources and schedule. Resources and schedule information will drive VV&A scope, which will have a major impact on technique tailoring.
- Personal Bias
  - Stakeholders and/or VV&A personnel involved in executing MURM may have preferences for and/or biases against various V&V techniques. Technique selection guidance from the M&S User and other stakeholders can also impact technique selection.

The following guidance options should be considered when selecting V&V techniques:

- Require V&V technique to satisfy a specified minimum level of information quality (techniques producing higher quality of information are acceptable).
  - **Expected Impact**: Such techniques may require expertise that V&V personnel in the project do not possess. Training of personnel may be required, or personnel with appropriate expertise obtained for the project, or both.

- Complete V&V assessment by a specified time or perform the V&V without all desired or required information.
  - **Expected Impact**: Techniques compatible with the time or information available may not have the desired information quality. Appropriate authority will have to modify the required information quality or accept a delay in the V&V assessment until appropriate information is available or a technique capable of producing the desired information quality can be performed.
- Give preference to V&V techniques that can be used for other capabilities and M&S aspects as well as for this capability.
  - **Expected Impact**: While providing information whose quality at least satisfies the required quality level, these techniques may not be the ones that can produce the best quality information or be performed most economically or rapidly.
- Give preference to the most economical and/or fastest V&V techniques that satisfy information quality requirements.
  - **Expected Impact**: This is ideal guidance if such techniques are within the expertise of available personnel. However, this guidance does not provide a hedge if changes occur which require higher levels of information quality. Because the information quality is only what was required initially, it may be necessary to redo the V&V to produce higher quality information. Original use of techniques producing better quality information might preclude the need for such rework of the V&V.
- Use the V&V techniques that produce the best quality information.
  - **Expected Impact**: This guidance may cause V&V to be more expensive and/or take longer than might be done with lower quality information.

2.2.3.5 V&V Techniques Catalog

The purpose of the V&V Techniques Catalog is to provide the details that a V&V practitioner would need to know to select the V&V techniques that best meet the needs and constraints of the V&V effort. The goal of the catalog is to serve as a comprehensive repository of techniques to facilitate the VV&A planning process.

The catalog builds on the information published in reference (d) and the DoD VV&A Recommended Practices Guide [Reference (f)]. The template used in the catalog provides a consistent way to describe the techniques and provides the details necessary to select the best technique among several choices for a given VV&A project.

See Appendix D for more detailed information about the V&V Techniques Catalog.

2.2.4 Computing p(Causes)

Once Factor C_1 [information about the clarity of the intended use], Factor C_2 [information about the importance of capability] and Factor C_3 [information about the recommendation to employ or not to employ M&S Results] (See Sections 2.2.2, 2.2.3.2, 2.2.3.3, and 2.2.3.4) are defined, then p(Causes) can be computed. See Appendix F for more detailed information about computing p(Causes).

2.2.5 V&V Implementation

Once the initial V&V Plan is developed and approved, the next step is to implement the plan that has been tailored based on the available information and available resources.

It is important to remember that change happens in the dynamic execution of programs and plans (e.g., requirements could change, schedule could change, resources could change, etc.). Just as program managers continually manage program risk, M&S Use Risk should be managed throughout the implementation of the V&V Plan. As program managers make decisions that could affect M&S Use Risk, VV&A practitioners must reassess the impact those program decisions will have on the implementation of V&V Plans. Any resulting changes in capability priorities, activities or techniques selected would require a recalculation of M&S Use Risk. VV&A practitioners can then provide tailoring solutions based on recalculated M&S Use Risk back to the stakeholders for approval.

The assignment of M&S Capabilities Importance also could be refined as evidence is collected and more information is gained. When this happens, VV&A practitioners must reassess M&S Use Risk. Generally, the timeframe for agreeing on how to tailor V&V is very short and so it is important that the reassessed potential M&S Use Risk can be communicated clearly to decision-makers.

A risk is only a risk until it happens, and then it becomes an issue that must be addressed. V&V are implemented to avoid issues in making decisions to use M&S. V&V are risk mitigation processes for M&S. There are five risk handing strategies relevant to M&S Use Risk:

- Control: Take active steps to minimize M&S Use Risk through tailoring.
- Monitoring: Watch and periodically reevaluate M&S Use Risk for changes.
- Transfer: Reallocate M&S Requirements to lower M&S Use Risk.
- Avoidance: Change or lower M&S Requirements.
- Acceptance: Acknowledge the M&S Use Risk and M&S limitation, but take no action.

MURM enables VV&A practitioners to employ the control, monitoring, and transfer risk handling strategies and to communicate M&S Use Risk effectively to stakeholders. V&V tailoring directly addresses control. VV&A practitioners take active steps to minimize M&S Use Risk by prioritizing M&S Capabilities and selecting appropriate V&V activities, tasks, and techniques. VV&A practitioners must also periodically reassess the potential M&S Use Risk based on the inevitable changes that will occur during implementation. As more information is gained about M&S capabilities and limitations, the associated M&S Use Risk possibly can be reallocated or transferred to other M&S requirements.

The V&V Report [see Reference (g)] documents the results of implementing the V&V Plan, captures the tailoring changes made during implementation, and produces a map of the M&S capabilities and limitations to enable M&S Users to shape the use of the M&S to take advantage of its capabilities and avoid the pitfalls of its limitations. The V&V Report communicates the M&S Use Risk associated with each M&S Capability explored during implementation of the V&V Plan.

While MURM provides detailed information used to derive the confidence in the V&V evidence (see Figure 2-9) and ultimately the M&S Use Risk (see Appendix F3), effective communication of this information to the M&S User requires focus on the key issues at this phase of the process: V&V results and the confidence in the evidence that produced those results. Figure 2-9 provides an example of a summary scorecard that could be used to communicate the more detailed mathematical results as shown in Table F-1 and Table F-9. This communication mechanism indicates for each capability: (a) if V&V were performed for that capability; (b) if the required capability exists in the M&S; and (c) the confidence in the evidence.

	M&S Capability Importance Value	V&V Results Show How Acceptability Criteria Addressed	V&V Evidence Confidence Value
Capability #1			
Capability #2a			
Capability #2b			
Capability #3			
...			
Capability #N-1			
Capability #N			

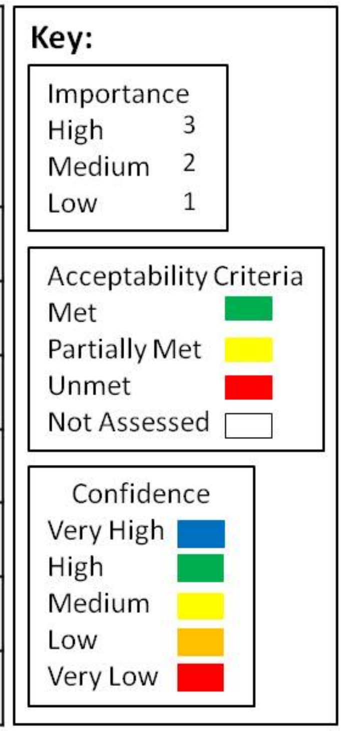


Figure 2-8: Example Communication Mechanism

The information documented in the V&V Report is assessed by the Accreditation Authority against the Acceptability Criteria, MOEs and MOPs. Accreditation Implementation, the Accreditation Decision, and the ultimate application of the M&S are addressed in the next subsection.

2.2.6 Accreditation Implementation, Accreditation Decision, and M&S Use

The Accreditation Agent responsible for producing the Accreditation Plan on behalf of the Accreditation Authority is also responsible for implementing that plan. Once the V&V Report and other accreditation information have been accumulated, they are assessed against the Acceptability Criteria and the associated MOEs and MOPs. The effects on the M&S Use are also updated as changes might have occurred that affect M&S Impact and M&S Reliance.

The summary scorecard in Figure 2-9 is updated with the results of the accreditation assessment on each M&S Capability and recommendations for accrediting the M&S for its ultimate use are produced and communicated to the Accreditation Authority.

The results of Accreditation Implementation are reported in the Accreditation Report [see Reference (g)]. Acting on behalf of the M&S User, the Accreditation Authority uses that information to make the decisions to accredit the M&S for use.

Based on information in the Accreditation Report and the recommendations for accreditation, decisions based on each M&S Capability are made and documented in an Accreditation Decision Letter. Several decision options are available to the decision-maker:

- Full Accreditation: The M&S Capability produces results that are sufficiently credible to support the application.
- Limited/Conditional: Constraints are placed on how the M&S Capability can be used.
  - Accreditation: Based upon the evidence assessed, the need for additional information to be provided, or modifications required to the M&S.
  - Non-accreditation: Results of the assessment show that the M&S Capability is not fit to support the application.

The tendency to aggregate M&S Use Risk for the individual M&S Capabilities should be avoided. Each capability is not equal and aggregation of M&S Use Risk could provide invalid M&S Use Risk information. Further research into the mathematical logic behind aggregation is needed.

Results of applying the M&S Capabilities should be captured in whatever reporting products are available (e.g., studies, test reports, etc.). These products are not only valuable for historical purposes, but for future uses, too. Additionally, relevant error information discovered during use should be provided back to the M&S Proponent and M&S Developer to address in future versions of the M&S.

While these reports provide the detailed findings of both the VV&A processes and the implementation of MURM, visualization tools (e.g., histograms) can be utilized to provide a high-level overview of where M&S Use Risk exists across the M&S requirements. Figure 2-10 shows the results for an M&S with 15 required capabilities and a confidence threshold for the V&V evidence set at medium or higher. The priority of the capability is designated as a number in the range 1—3, the results of the V&V efforts for a capability is designated by color, and the computed confidence of the V&V evidence is designated by the height or depth of the bar.

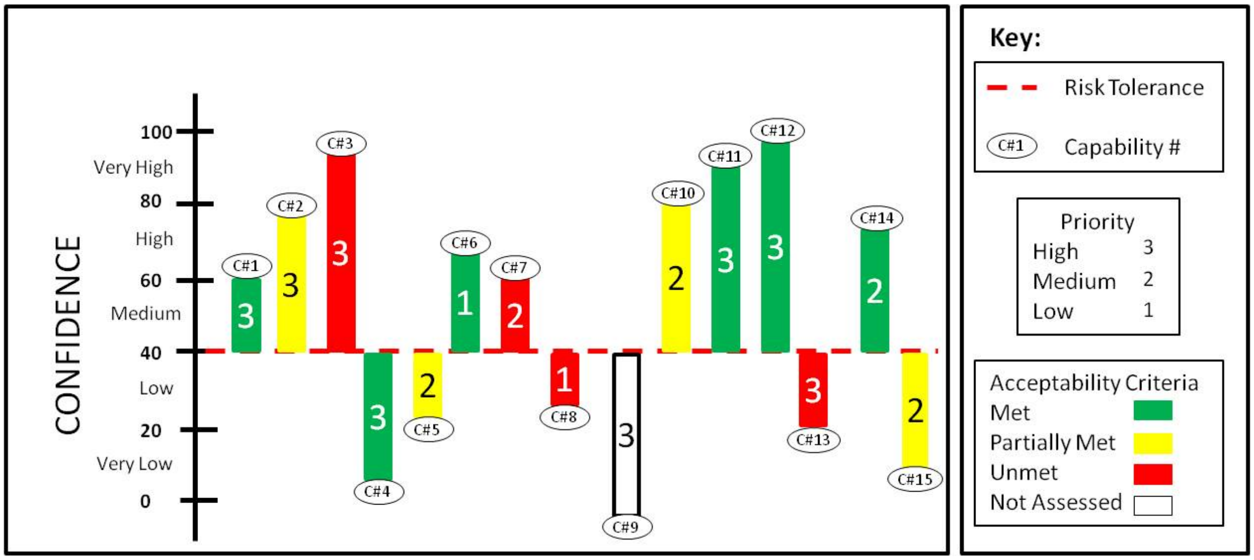


Figure 2-9: Visualization of M&S Use Risk

For example, C#4 is a high priority capability that met the acceptability criteria (as shown by the associated green bar) but with no confidence in the accuracy of that result (as shown by C4 bar falling below the tolerable risk level.)

This page intentionally left blank.

3. SUMMARY AND RECOMMENDATIONS

3.1 SUMMARY

The MURM exploits several existing concepts to create a practical, versatile, and disciplined approach to VV&A planning and execution. It focuses the V&V effort upon those M&S Capabilities that drive M&S Use Risks and incorporates techniques for M&S Use Risk assessment and communication as integral parts of the methodology. The disciplined execution of MURM will produce V&V evidence that describes M&S capabilities and limitations and the risks of encountering those limitations in the intended use. This information can help M&S Users direct their use of the M&S to avoid encountering the consequences of those risks.

Steps taken to derive MURM began with the translation of the M&S Use Risk definition into a mathematical logic expression. This allowed M&S Use Risk to be evaluated based on the probabilities of its elements (i.e., the causes and effects).

A detailed analysis of the three causes (i.e., clarity, importance, and confidence) that contribute to the calculation of M&S Use Risk was performed and state tables were derived. A scorecard was developed to keep track of the relevant information that determines the probability of the causes, the probability of the effects, and the resulting M&S Use Risk, and to support trade-off decisions.

Additionally, the MURM report provides information describing ancillary sources of information important to implementing MURM:

- V&V Process Maturity Model
- V&V Composite Model
- V&V Techniques Catalog

3.2 RECOMMENDATIONS

M&S Users should adopt MURM to support implementation of VV&A.

Improvements to MURM should focus on enhancing the methodology through the application of lessons learned and feedback from organizations using MURM.

MURM should be developed into an automated tool that provides a user-friendly interface and automated processing of the mathematical logic calculations.

MURM provides an assessment of M&S Use Risk at the M&S Capability level. Further research into the mathematical expression of aggregated M&S Use Risk is needed.

Additional effort is required to complete the VCM and to customize the tailoring approach for Accreditation.

FY10 Use Case applications of the MURM concepts have shown the value of the VPMM. Further refinement the VPMM as a guidance product would provide significant value.

APPENDIX A: REFERENCES

(a) Department of Defense Modeling and Simulation Coordination Office, 1901 N. Beauregard Street, Alexandria, Virginia 22311-1705, http://msco.mil/.

(b) The Johns Hopkins University Applied Physics Laboratory. The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Literature Survey and Methods Bibliography, JHU/APL NSAD-R-2009-207. October 2009.

(c) The Johns Hopkins University Applied Physics Laboratory. The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Assessment of Applicable Methods, JHU/APL NSAD-R-2010-001, January 2010.

(d) The Johns Hopkins University Applied Physics Laboratory. An Approach for Realizing a Risk-Based VV&A (RBA) Methodology, JHU/APL NSAD-R-2010-020, March 2010.

(e) The Johns Hopkins University Applied Physics Laboratory. An Approach for Realizing a Risk-Based VV&A (RBA) Methodology – Review and Comments Summary, JHU/APL NSAD-R-2011-018, March 2011.

(f) Department of Defense Verification, Validation, and Accreditation Recommended Practices Guide, Build 3.0, September 2006, http://vva.msco.mil/.

(g) Modeling and Simulation Coordination Office. MIL-STD-3022, Department of Defense Standard Practice Verification, Validation, and Accreditation for Models and Simulations, Defense Standardization Program, January 2008, https://assist.daps.dla.mil/quicksearch/.

(h) The Johns Hopkins University Applied Physics Laboratory. Verification, Validation and Accreditation Activity and Task Tailoring, 2011.

(i) Department of Defense Modeling and Simulation Coordination Office. *Draft Department of Defense Standard Practice for the Evaluation of the Maturity of Model and Simulation Validation Processes*, September 2008.

(j) Youngblood, S.M., & Harmon, S.Y., "A Proposed Model for Simulation Validation Process Maturity," *The Journal of Defense Modeling and Simulation Applications, Methodology, Technology* 2(4): 179-190, October 2005.

(k) The Johns Hopkins University Applied Physics Laboratory. Final Report of Modeling and Simulation Group 054 / Task Group 037, JHU/APL NSAD-R-2010-022, March 2010.

(l) Hubbard, D. W., *The Failure of Risk Management*, John Wiley & Sons, Inc., Hoboken, NJ (79-93), 2009.

(m) Devore, J.L., *Probability and Statistics*, Sixth Ed., Thomson, Brooks/Cole, Belmont, CA, (63), 2004.

(n) Jaynes, E. T., *Probability Theory The Logic of Science*, Cambridge University Press, First published 2003, Reprinted with corrections 2004.

(o) Jaynes, E. T., "Where Do We Stand On Maximum Entropy?" Maximum Entropy Formalism Conference, Massachusetts Institute of Technology, May 1978.

(p) Jaynes, E. T. "Where Do We Go From Here," *Maximum Entropy and Bayesian Technology*, May 2-4, 1978.

(q) Jaynes, E. T. "Where Do We Go From Here," In *Maximum Entropy and Bayesian Methods in Inverse Problems*, C. Ray Smith and W. T. Grandy, Jr. (Eds.), D. Reidel Publishing Company (21-58), 1985.

(r) IEEE Std 1012™-2004, IEEE Standard for Software Verification and Validation, IEEE Computer Society, Copyright © 2005 by the Institute of Electrical and Electronics Engineers, Inc., April 2005.

(s) Department of Defense. Department of Defense Instruction 5000.61 DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A), December 2009, http://www.dtic.mil/whs/directives/corres/pdf/500061p.pdf/.

APPENDIX B: V&V ACTIVITY AND TASK TAILORING

The primary purpose for tailoring the V&V process is to create a V&V Plan that identifies the activities and tasks that will be performed to develop the V&V evidence to support the application of a specific model or simulation to a specific intended use. Tailoring may also modify an existing V&V Plan after the V&V process has begun. This approach to V&V tailoring seeks to produce effective and efficient V&V Plans so that V&V information can be obtained with adequate scope and quality to support acceptable use risk levels within schedule and resource constraints. If that were not possible then this approach would produce a V&V Plan to obtain the best information within available schedule and resource constraints. The proposed V&V tailoring approach essentially customizes the activities and tasks in the V&V Composite Model (VCM) (see Appendix E).

The following discussion only deals with the selection of V&V activities and tasks. The choice of the techniques through which to execute the selected tasks is the topic of future research. Therefore, this approach depends upon the following assumption:

Assumption 1: The VCM defines all of the activities and tasks that could be executed for any model or simulation for most intended uses.[3]

In other words, the proposed V&V approach chooses those parts that apply to the specifics of the situation. In actuality, the VCM partitions the executable components of the V&V process into phases (e.g., plan the V&V effort, verify and validate the M&S conceptual model). It then divides each V&V phase into activities (e.g., verify the M&S development products) and these activities into separate tasks. This hierarchical partitioning of V&V tasks permits V&V planning to occur hierarchically by first selecting the V&V phases relevant to the situation then the activities associated with each selected phase then the tasks associated with each selected activity.

The VCM breaks the V&V effort into eight phases. The hierarchical partitioning of V&V phases, activities and tasks suggests that a V&V Plan can be

[3] The VCM may not address the needs of all possible models or simulations, particularly those for which special approaches may be needed (e.g., an anechoic chamber for a hardware-in-the-loop missile M&S that provides real radio frequency (RF) signals to a seeker of a missile where the V&V effort must take into account facility factors such as temperature, humidity, vibration and alignment controls).

considered as a story that contains a beginning, middle and end. This analogy leads to two more assumptions that underlie the proposed V&V tailoring approach:

Assumption 2: Every V&V effort begins by producing a V&V Plan in some form.

Assumption 3: Every V&V effort ends by producing a V&V Report in some form.

These two assumptions are consistent with the view that every V&V effort should produce two primary deliverables: a V&V Plan and a V&V Report. Therefore, the beginning of every V&V Plan story describes the activities and tasks that will be executed to produce the V&V Plan and the end of every V&V Plan story describes the activities and tasks that will be executed to integrate the collected V&V evidence into the final V&V Report. The middle of the story describes how the V&V evidence will be collected. The implication of a V&V Plan requiring a middle that describes how the V&V evidence will be collected suggests another assumption.

Assumption 4: Every V&V effort performs one or more activities that collect V&V evidence (i.e., Apply Historical Evidence, V&V the Conceptual Model, Perform Supplemental Verification & Apply the Verification Products to Validation, Validate the M&S Results).

This model of a V&V Plan reduces the selection of eight phases to choosing from among six phases. To follow the story analogy further, the choices of phases to be executed define the paragraphs in the story. Thus, each story will begin with a planning paragraph, end with a reporting paragraph and contain paragraphs that describe the collection of the V&V evidence. This model and the underlying VCM account for those situations where a V&V Plan may evolve from an initial plan, which may be very abstract with little detail, into a final one, that will contain considerably more resolution. This model also accounts for iterative V&V situations that produce a number of V&V Reports during the M&S lifecycle.

The choice of activities through which to collect V&V evidence can be further reduced through the following assumption:

Assumption 5: Every V&V effort will include the V&V of data or knowledge sets as part of the activities for collecting the V&V evidence.

Therefore, the middle of every V&V Plan story will include a paragraph to represent the activity for verifying and validating the available data or knowledge sets. This reduces the choices of phases for collecting the V&V evidence to five:

1. Apply the relevant historical information.
2. Verify and validate the M&S conceptual model.
3. Perform supplemental verification.
4. Apply the verification products to validation.
5. Validate M&S Results.

Phases 3 and 4 both deal with the verification of the development products. Phase 3 describes those verification activities that the V&V Team would perform and Phase 5 leverages the verification evidence that both the V&V Team and developers collect. These two phases can be grouped logically together into a single activity called verify the development products. This grouping reduces the number of somewhat independent choices of activities for collecting V&V evidence to the following four:

1. Apply the relevant historical information (history application)
2. Verify and validate the M&S conceptual model (CM V&V)
3. Verify the development products (DP verification)
4. Validate M&S Results (results validation)

These four phases can be combined into the following practical combinations:

1. CM V&V
2. DP verification
3. history application or (CM V&V + DP verification)
4. history application + CM V&V
5. history application + DP verification
6. history application + CM V&V + DP verification
7. results validation
8. CM V&V + results validation
9. history application + results validation) or (DP verification + results validation
10. history application + CM V&V + results validation
11. history application + DP verification + results validation

12. CM V&V + DP verification + results validation

13. history application + CM V&V + DP verification + results validation

Since order is important, 24 possible combinations can be constructed from four independent elements. However, the inherent causality of the M&S development lifecycle makes some of these orderings infeasible and others impractical. For example, development of the M&S conceptual model occurs in the first phases of the design process whereas the detailed design, source code production and executable production occur afterward. This natural ordering of the M&S development lifecycle makes the M&S conceptual model available before the rest of the development products. As a result, the combination CM V&V and DP verification is practical where the combination DP verification and CM V&V is not. Finally, the opportunity to benefit from prior information eliminates still other combinations. For example, results validation should follow all of the other activities for collecting V&V evidence because it can only be performed toward the end of the M&S development lifecycle and can benefit from the evidence produced by those preceding activities. In addition, results validation must follow V&V of data and knowledge sets since results validation testing cannot be performed without valid data and knowledge sets. This includes those situations where the conceptual model is used to gain insight into the predictive behavior of the model or simulation in those areas where referent data is either unavailable or sparse.

The list above is roughly in order of improving evidence (where the criteria are coverage and certainty or confidence). Two sets of activities have been grouped in this list because their value seems approximately equal. For example, CM V&V produces the least valuable evidence because M&S conceptual models tend to be abstract descriptions of M&S functionality that are produced many steps before the actual executable M&S is realized. This is especially true for models or simulations that can evolve with use (e.g., adaptive M&S). Each of these steps introduces uncertainty into the V&V evidence from the CM V&V activity. DP verification produces more relevant information about the M&S actual functionality but the current theory and tools only guarantee incomplete evidence at best. Those options that combine evidence from multiple sources can produce stronger information than those that rely upon a single source. The exception to this observation is results validation since it provides the most direct evidence of the actual functionality of a model or simulation.

This list assumes that all of the information needed to perform the activities is available to the V&V effort. It gets simpler if one or more of the individual sources of evidence are not available (e.g., historical information). This list further assumes that all of the information needed is available and that each individual activity provides evidence of value in the order shown (i.e., CM V&V < DP verification < history application < results validation).

The preceding paragraphs identify the first level of V&V tailoring decisions. The remainder of this document describes the decisions associated with tailoring the activities and tasks from the VCM. Again, following the story analogy, the activities represent the sentences in the activity paragraphs and the individual tasks represent the words used to compose the activity sentences. Composing activity sentences relies upon the following assumptions:

Assumption 6: Every V&V task takes information as input, either explicitly described or assumed.

Assumption 7: Every V&V task produces information as output.

Assumption 8: A V&V task cannot be executed unless the tasks that produce the information that it needs as input are executed beforehand. This suggests that the sentences that describe the V&V effort are temporally organized.

These assumptions define the general causality that determines the order and possible combinations of activity sentences from the constituent tasks. Complete and detailed tailoring descriptions for all the V&V activities and tasks can be found in a separately published document [Reference (h)].

Associated activities and tasks for Accreditation were not included in the VCM (see Appendix E). Therefore, additional effort is required to complete the VCM and to customize the tailoring approach for Accreditation.

This page intentionally left blank.

.

APPENDIX C: SUMMARY OF THE V&V PROCESS MATURITY MODEL

The Verification and Validation (V&V) Process Maturity Model (VPMM) describes a practice for evaluating the maturity of V&V processes. This practice recognizes V&V as processes that generate information as their sole products.

The VPMM substitutes process objectivity and the properties of quality information for the difficult-to-measure qualities of truth. The properties of information quality are defined by the completeness and correctness of the information and by the confidence the information producer has that the information is complete and correct enough to serve an intended use.

V&V processes take acceptability criteria, the referent, any available historical information, the M&S conceptual model, verification results of the intermediate development products, information about the data and knowledge sets needed to support the intended use, and the M&S Results as input and produce evidence assessing the M&S validity as output.

The quality of the input information ultimately limits the quality of the evidence that the V&V processes produce. This practice for evaluating process maturity is structured by maturity and defined by the quality of the input information, the quality of its information products, and the objectivity of those process components that contribute to those products. Table C-1 provides information about each level.

Table C-1: VPMM Levels Descriptions

Maturity Level	Acceptability Criteria	Referent	Conceptual Model	Development Products	M&S Results
1	represented by Subject Matter Expert (SME) opinion	represented by SME opinion	if one exists, validated by SME review	verified only enough to support development	validated by SMEs observing simulation results
2	determined from user statements in terms of entities represented, their attributes and the dependencies between them	represented solely by SME opinion	validated against the validation criteria by the SME	verified against the conceptual model inventory	validated by SME against the validation criteria
3	determined from user statements in terms required for Level 2 plus the attribute ranges, domains and errors	derived from a single source with estimates of errors	validated by objective party from validation criteria and referent	verified against the conceptual model	evaluated by objective party from validation criteria and referent
4	determined from user statements in terms of entities, attributes, ranges, domains, errors and confidences	sampled from multiple independent sources and correlated statistically with estimates of uncertainties (i.e., errors and confidences)	validated by objective party from validation criteria and referent; analyzed to suggest results sampling space and estimate the confidence associated with that sampling	verified against the conceptual model; provides information to guide results sampling and estimate the confidence associated with that sampling	sampled from guidance developed from CM and verification results analysis; validated by objective party from validation criteria and referent

Table C-1: VPMM Levels Descriptions (continued)

Maturity Level	Acceptability Criteria	Referent	Conceptual Model	Development Products	M&S Results
5	formally derived from user statements using causality arguments	formally derived from multiple independent sources and characterized statistically with estimates of uncertainties	formally stated and validated through proof against the validation criteria and referent; analyzed to define results validation sample space	verified against conceptual model and used to define results validation sample space and the confidence associated with that sampling	sampled from guidance defined from CM and verification results analysis; validated through proof against the validation criteria and referent

ANALYSIS OF VPMM USE CHARACTERIZATION

Mathematical analysis of the information in Table C-1 produces 5 logical atoms, with $2^5 = 32$ combinations of which 10 are in the solution space and 22 are outside of the solution space.

Maximum Information Entropy (MIE) Weightings and Subsequent Percentile Distribution

Logical Atoms
v = SME judgments
w = representation inventory and primary use risks for inventory groups
x = tolerable error characterizations and use risk estimates for errors and bounds)
y = tolerable uncertainties and sensitivities
z= mathematical derivation description

In Table C-2, positive atoms are bolded. An explanation of the mathematical notation can be found in Appendix G.

Table C-2: VPMM Mathematical Analysis

	Logical Sentence	URS	No. of Logical Sentences	Cumu-lative	Percen-tile	Quin-tile	Color	Word
A	~v^~w^~x^~y^~z	1	1	1	0	1st	Blue	Very Low
B	**v**^~w^~x^~y^~z	3	1	2	11.1	1st	Blue	Very Low
C	(**v**,~v)^**w**^~x^~y^~z	6	2	4	33.3	2nd	Green	Low
D	(**v**,~v)^**w**^**x**^~y^~z	10	2	6	55.6	3rd	Yellow	Medium
E	(**v**,~v)^**w**^**x**^**y**^~z	14	2	8	77.8	4th	Orange	High
F	(**v**,~v)^**w**^**x**^**y**^**z**	18	2	10	100	5th	Red	Very High

Figure C-1 depicts the analysis in Table C-2. The colors used relate to the colors in Figures 2-2 and 2-3 in Section 2.2 of the main body and Figure C-2.

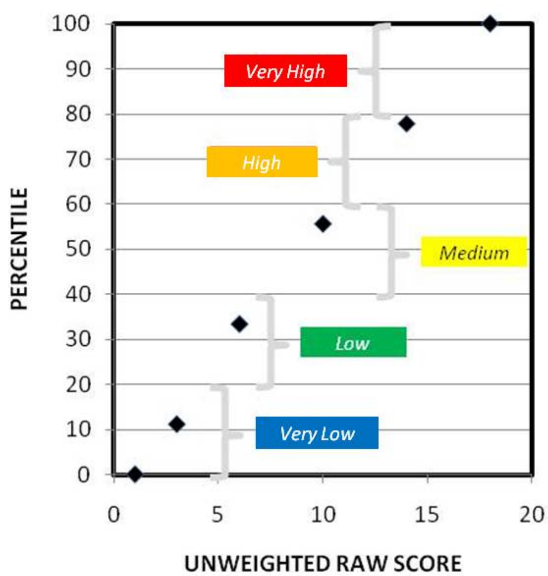


Figure C-1: VPMM Unweighted Raw Score

The coloring of the VPMM in Figure C-1 illustrates the approximate relationship between the V&V process maturity levels. To focus VV&A efforts, decision-makers must make tradeoffs between the costs of knowing that the M&S to be used is acceptable for the intended use against the risk of assuming that they know. As objectivity increases in the collection, evaluation, and assessment of V&V evidence, the relative cost of V&V increases, too.

MURM		V&V Process Maturity Level				
		Very Low	Low	Medium	High	Very High
Clarity of Intended Use (Factor C_1)	Acceptability Criteria	represented by SME opinion	determined from user needs in terms of entities represented, attributes and dependencies	Low + plus attribute ranges, domains & errors	Medium + entities and confidences	**High + formerly derived using causality**
	Validation Referent	represented by SME opinion	represented by SME opinion	derived from a single source with estimates of errors	Medium + multiple independent sources and correlated statistically	**High + formerly derived and characterized statistically**
Confidence in V&V Evidence (Factor C_3)	Conceptual Model	if one exists, validated by SME review	validated against the acceptability criteria by SME	validated by objective party from acceptability criteria & referent	Medium + analysis of results sampling space & estimate of sampling confidence	**High + formerly validated through proof**
	Development Products	verified only enough to support development	verified against the conceptual model	verified against the conceptual model	Medium + guide results sampling & estimate of sampling confidence	**High + verified against conceptual model and used to define confidence**
	M&S Results	validated by SME observation of M&S results	validated by SME against the acceptability criteria	evaluated by objective party from acceptability criteria & referent	Medium + sampled from guidance developed from CM & verification results analysis	**High + sampled and validated through proof**

Relative Cost of Increasing Objectivity In V&V Processes

$	$$	$$$	$$$$	$$$$$

Figure C-2: Validation Process Maturity Model

Figure C-2 depicts the VPMM as defined in Table C-1 overlaid with the colors indicating the relative V&V rigor. As defined in Figure 2-3, as rigor increases so does cost. The Very Low maturity level consists entirely of face validation where the validation evidence depends entirely upon subjective sources of requirements, referent, and validity judgments. The Low maturity level improves the objectivity and quality of the representational requirements. The Medium and High maturity levels progressively improve the objectivity of the conceptual modeling and results validation component processes and the quality of the evidence they produce. In the Very High maturity level, the validation process rigorously transforms informal user need statements into formal acceptability criteria, then applies mathematical techniques to prove conceptual model and M&S Results validity. In the VPMM, the objectivity and quality of the validation evidence increases as the level of process maturity increases.

Deciding to apply M&S Results with no V&V evidence leaves those M&S users in the most uncertain situation with little knowledge about the M&S capabilities and limitations. This forces them to assume all of the risks of using the M&S Results. Providing V&V evidence based upon expert opinion gives them information but information that lacks reproducibility and contains all of the uncertainties associated with expert opinions. The M&S users might have information about M&S capabilities and limitations, but it is uncertain at best and the degree of uncertainty is largely unknown; therefore, the risks of using that information are largely uncharacterized. As the maturity level of the V&V processes increases, the M&S user gets more and higher quality information about the M&S capabilities and limitations with decreasing uncertainty and better characterization of that uncertainty that exists in the V&V evidence.

Information about the VPMM is available through M&S CO in an unpublished *Draft Department of Defense Standard Practice for the Evaluation of the Maturity of Model and Simulation Validation Processes* [Reference (i)]. Additionally, an article published in the *Journal of Defense Modeling and Simulation* provides information about the VPMM [Reference (j)].

APPENDIX D: SUMMARY OF THE V&V TECHNIQUES CATALOG

The V&V Techniques Catalog is one of the cornerstone components of MURM. The purpose of the techniques catalog is to provide the details that a VV&A practitioner would need to know to select the techniques that best meet the needs and constraints of the V&V effort. The catalog builds on the methods assessment reports and methods taxonomy described in Appendix C and Appendix D of Reference (c), as well as the techniques included in the DoD VV&A Recommended Practices Guide (RPG) [Reference (f)]. The goal for the catalog is to serve as a comprehensive repository of techniques to facilitate the V&V planning process.

The MURM effort began by surveying the relevant technical literature to identify methods that could support a risk-based VV&A methodology. This effort continued by assessing several of the methods discovered in the literature survey. The method assessment combined the assessed methods with the V&V technique taxonomy from the DoD VV&A RPG to create a collection of over 200 methods that could support MURM. This combined collection was then organized into an integrated taxonomy of methods. The individual method assessment reports also combine with the descriptions from the VV&A RPG to form an initial catalog.

Figure D-1 illustrates the five basic tasks that are essential to V&V planning and shows where the V&V Techniques Catalog is accessed when planning to verify and validate M&S. Each technique has a specific purpose, needs particular information as input, requires certain knowledge, skills, or tools to apply, and outputs specific information. Based on the knowledge, skills, and/or tools required to apply the technique, the V&V resources needed may differ. The techniques run the gamut across the test, analysis, demonstration, and inspection categories.

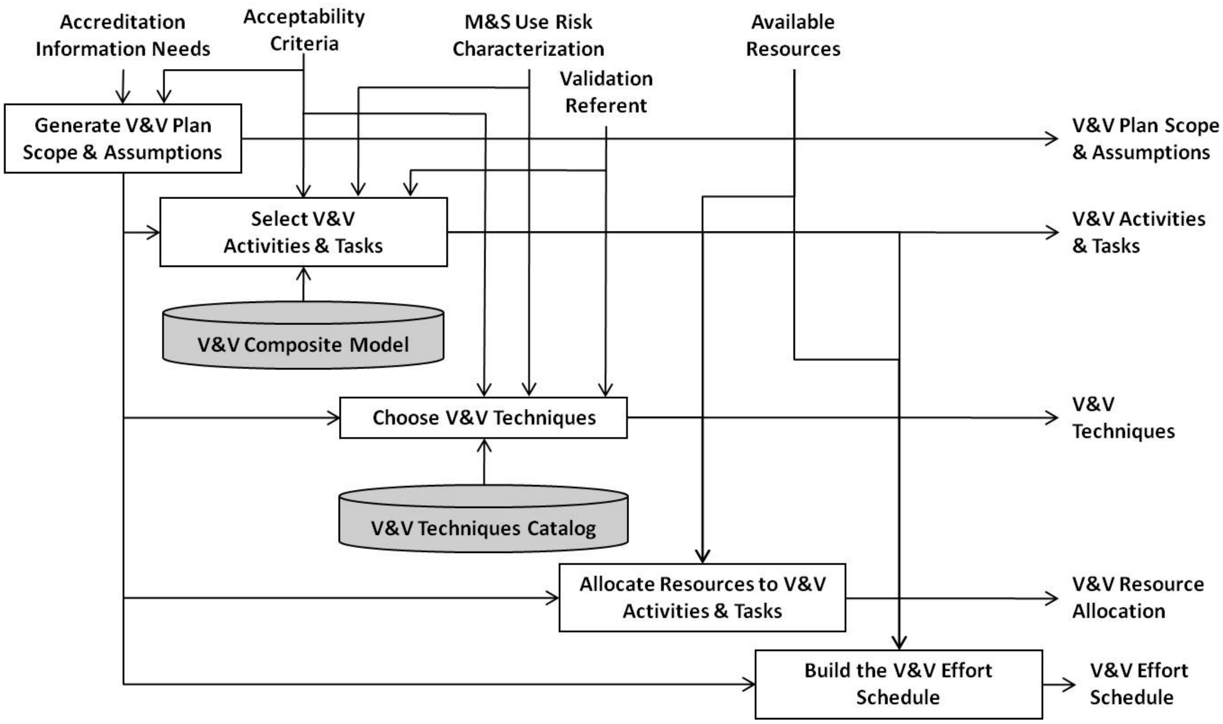


Figure D-1: V&V Techniques Catalog Function in MURM

The V&V Techniques Catalog itemizes all of the techniques that could be applied to executing the tasks from the VCM and identifies the minimum resources needed to apply those techniques. Following is an explanation of the information captured for the Bayesian Risk Analysis technique. Each element of information in the standard template is explained below. Additionally, Figure D-2 provides a sample of the actual template for Bayesian Risk Analysis.

- Technique Name: Each technique template begins with name of the technique.
- Technique Category: This field shows the relationship of the technique to its subcategory (Risk Analyses), category (Analyses), and phase (Information Collection).
- Mapping to VCM: The mapping relates the technique to a particular part of the VCM (see Appendix E). In this example, Bayesian Risk Analysis falls under the process named "Develop the recommendations for use" and is applied to the sub-process "Analyze the V&V Evidence." This component of the template fulfills the objective to relate the technique to the VCM. The mapping comes from Table 4-11 in Reference (c).

- Keyword Search: This field contains some words that might be used to search for information on the technique.
- Technique Description: This field provides a high-level overview of the technique. The descriptions come from References (c) and (f).
- Technique Process: This field is an overview of the steps to accomplish the technique in sufficient detail to determine if the technique meets the constraints of the VV&A effort in terms of human skills and levels, schedule, budget, and risk.
- Technique Assessment: This field comes from Reference (c). The assessments were provided by the SMEs who researched the techniques for the Assessment Report [Reference (c)].
- Technique Influences: Human resource factors that influence the selection of a technique are described in terms of the skills, skill levels, and education levels required to execute the technique selected successfully.
- Level of Effort is a red (High), yellow (Medium) or green (Low) indicator of the relative amount of effort required to execute the technique. This may be influenced by equipment requirements, time requirements, or cost.
- Technique Information Requirements: This field defines the required data input to execute the technique. It may also outline necessary pre-conditions.
- Technique Output: This field defines the format, content, and measures associated with the results of executing the technique.

The catalog is in the initial stages of development and requires enhancement, refinement, and additional research before release for technical review and comment.

Technique Name: Bayesian Risk Analysis

Technique Category

Information Collection → Analyses → Risk Analysis

Mapping to V&V Composite Model

Develop the recommendations for use in "Analyze the V&V Evidence"

Keyword Search

- Risk Management
- Risk Assessment
- Uncertainty Analysis
- Risk-Based VV&A
- Bayesian

Technique Description. Bayesian Risk Analysis combines Bayesian updating with risk analysis. A prior risk assessment may be done with prior assumptions on the uncertain variables. As more data become available, Bayesian analysis is used to update the distributions of the uncertain variables and the risk estimate. Since systems have multiple components, multiple variables that could be updated, and multiple risk events, a Bayes network approach can be used to systematically combine all the variables, models, and available data.

Technique Process [Describe how to accomplish the technique here]

Technique Assessment. The Bayesian approach is expected to be highly effective for M&S applications. This type of analysis can take into account data of different types, such as expert opinion and experimental results, at different levels of the M&S activity. The resources required for applying this method are related to data collection, which include model outputs, laboratory experimental data, field performance data, as well as expert opinion. The method applies to Risk Analysis and Risk Tracking modules of the DoD Risk Management Process. Recommended Applicability—The method looks very relevant and applicable.

Technique Influences

Human Resource: (required skills and skill levels education, numbers) – need to figure out

Level of Effort	High	Medium	Low

Technique Information Requirements (enter text here)

Technique Output (enter text here)

Technique Maturity

Maturity Level:			
Theoretical Base:	Solid	Questionable	Unknown
Application:	Extensive	Limited	Unknown
Guidance:	Extensive	Limited	Unknown

References

- Jiang, X., and Mahadevan, S., "Bayesian Risk-Based Decision Method for Model Validation under Uncertainty," *Reliability Engineering and System Safety*, 92(6):707–718, June 2007
- Wilson. A. G.. and Huzurbazar. A. V.. "Bavesian Networks for Multi-level Svstem

Figure D-2: Bayesian Risk Analysis Technique Template

APPENDIX E: SUMMARY OF THE V&V COMPOSITE MODEL

The V&V Composite Model (VCM) is a product of the Research and Technology Organisation of the North Atlantic Treaty Organisation Modeling and Simulation Group MSG-054/TG-037. This report documents the V&V portions of the VCM:

- Final Report of Modeling and Simulation Group 054 / Task Group 037, March 2010, Chapter 3 – Description of V&V Composite Model, JHU/APL NSAD-R-2010-022 [Reference (k)]

Chapter 3 of the report describes the VCM. Information taken from that report appears below. The report is available through M&S CO [Reference (a)]. The report does not address the range of phases, activities, and tasks for Accreditation.

E-1. APPROACH TO V&V TAILORING

Enabling the tailoring of V&V processes begins with describing the range of phases, activities, and tasks that possibly could be performed when implementing V&V processes. The efforts of MSG-054/TG-037 resulted in the description of the phases, activities, and tasks for the V&V processes only. The phases, activities, and tasks for the accreditation process were not completed as a result of these efforts.

The VCM describes the components of the processes (i.e., phases, activities, and tasks) from which to select to match the risk and resource constraints of the V&V efforts while still adhering to relevant policies, standards, and guidance. The VCM is a superset of the possible activities and the context in which those activities can be tailored into working V&V processes.

E-2. COMPONENTS OF TAILORING GUIDANCE

Using risk as an effective tailoring mechanism for VV&A processes depends on understanding the factors that drive the risk, as well as the VV&A activities and tasks that can and should be implemented to address the use risks. This model of the V&V processes defines the tasks that may be tailored to best suit the limits of acceptable use risk and optimize the resources available for VV&A.

As the complexity of the M&S and the associated V&V processes increase, while budgets and schedules decrease, it is likely that not all V&V activities and tasks will be able to be performed. Additionally, when V&V is performed on legacy M&S, legacy developmental products may not be available to review, analyze, and test. To avoid becoming overwhelmed by the tasks described, to utilize resources efficiently, and to mitigate risk effectively, tailoring of the V&V processes is needed. Tailoring guidance helps determine the activities required for different V&V implementations.

Tailoring guidance is embedded within the descriptions of the activities in the VCM. This guidance helps the V&V practitioner determine what tasks to omit and to understand the corresponding risk.

E-3. OVERVIEW OF THE V&V COMPOSITE MODEL

The VCM comprises phases, activities, and tasks that decompose the general V&V processes. The highest levels are phases and there are eight phases in the VCM.

- *PHASE 1: Plan the V&V Effort.* The V&V processes are planned to address the acceptability criteria (developed during planning for accreditation) and the requirements for using the M&S. The V&V practitioner executes and evolves the V&V Plan throughout the V&V effort. This includes monitoring the V&V effort and adjusting the V&V Plan to reflect better any new information gained while executing the V&V processes. The V&V Plan should be approved by the individual responsible for ensuring the V&V processes are implemented.

- *PHASE 2: Apply Relevant Historical Information.* When a legacy M&S is to be applied for the same or a similar intended use for which it has been previously accredited, considerable historical information may exist upon which to base the new accreditation decision. In those instances, it is possible the entire V&V effort could focus upon the collection and analysis of historical information. When a legacy M&S is changed, when the simuland changes, or when the intended use changes, historical information may contribute to identifying the gaps in the additional V&V evidence that will be needed. Applying historical information begins by determining how closely the prior intended uses match the current intended use to determine what part of the history is relevant to the current problem. Historical information can come from prior V&V and testing activities, the records of prior uses, and the developer's accounts of the M&S capabilities and limitations. The V&V practitioner should analyze the relevant historical record to identify the factors that may constrain use. Then, a coherent picture of the M&S capabilities and limitations based upon the historical evidence can be assembled. If results validation will be performed, the V&V practitioner can use the historical evidence to identify possible output sampling areas to improve its efficiency.

- *PHASE 3: Verify and Validate the Conceptual Model.* The conceptual model serves as a bridge between the requirements for developing an M&S and the M&S design, providing the developer's interpretation of the requirements. The conceptual model is the documented theoretical approach to the design of the M&S. The conceptual model describes what the M&S is expected to do, to be, and what data and other elements are needed for a successful outcome. The V&V practitioner begins by characterizing its coverage of the intended use and inferring the M&S capabilities from the information that the conceptual model contains. Conceptual model verification involves checking it for internal consistency problems. After that, the V&V practitioner can evaluate the conceptual model against the acceptability criteria to determine its validity. The objective of performing conceptual model validation is to demonstrate that the M&S functional elements accurately and completely represent the M&S requirements and to identify where assumptions, limitations, or architectural structure impact the intended M&S use. If available, the V&V practitioner should also verify and, if needed, validate the use scenarios.

- *PHASE 4: Perform Supplemental Verification.* How much the V&V practitioner contributes to development product verification depends largely upon the verification activities performed and documented by the developer of the M&S. The V&V practitioner begins the supplemental verification by collecting the developer's development products and determining how much additional verification is needed to achieve the desired confidence. The types of verification tasks that can be performed depend largely upon the development products available. The V&V activities performed and documented by the developer against the requirements for building the M&S can be used to support the M&S validation activities performed by the V&V practitioner. The V&V practitioner should leverage as much of the developer's V&V evidence as possible and only perform what supplemental verification is needed to increase the confidence in and bolster the validation evidence.

- *PHASE 5: Apply The Verification Products to Validation.* During this phase, the V&V practitioner applies the collected verification information to develop evidence on the M&S validity, adds to the information to support the accreditation recommendations, estimates the coverage and uncertainties in the development product verification evidence, and employs the verification evidence to identify factors that may constrain use and candidate output sampling areas.

- *PHASE 6: Verify and Validate the Data and Knowledge Sets.* V&V is performed on the M&S. This phase focuses upon the data used to build the M&S, the data used as input into the M&S, and the data output by the M&S. Data and knowledge sets are distinguished by definition. Knowledge is a class of data that incorporates a knowledge representation scheme (e.g., production rules, semantic network, etc.), while data are representations of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means. Data and knowledge set V&V should be performed because of their independence from the M&S and because the M&S relies upon those sets to produce its output. The phase begins by identifying their sources and the pedigrees of those sources because in many cases the organizations responsible for producing the data differ from the organizations that either develop or use the M&S. The credibility of those data-producing organizations often weighs heavily in determining the credibility of the data itself. After establishing the pedigrees of the data and knowledge, the V&V practitioner should verify the internal consistency of the datasets; verify any transformations used to make the data accessible and meaningful to the M&S (e.g., units transformations, coordinate transformations); then validate the completeness and correctness of the datasets against the acceptability criteria. The data validation information can then be used to identify output sampling guidance and factors that may constrain M&S use.

- *PHASE 7: Validate the M&S Results.* During this phase, the V&V practitioner reviews, analyzes, and tests the M&S. Tests are run to produce output; the output is analyzed to determine capabilities; and the capabilities are reviewed against the acceptability criteria to infer validity. The V&V practitioner uses the developer's test results to tailor the results validation effort, if those results are available. Design of experiments techniques are applied to build the test cases used to sample the M&S output. The V&V practitioner executes the test cases, collects the output, analyzes the output, and documents an integrated description of the M&S capabilities and limitations from the collected V&V evidence.

- *PHASE 8: Integrate the V&V Evidence.* The preceding phases produce the evidence that this phase integrates into a description of the M&S validity to support the accreditation recommendations. The V&V practitioner examines the evidence produced by the preceding activities, forms a consistent picture of the M&S validity and estimates the confidence in that determination. The V&V practitioner identifies the factors that may constrain use and combines those constraints with the assessments of the M&S completeness and correctness for the intended use. Then, the V&V practitioner assembles the results of this integration with the evidence produced by the prior V&V activities into the V&V Report. The information in the V&V Report should present the evidence upon which an accreditation assessment can be made to base the accreditation recommendations. This phase may include any post-execution follow-up and archival activities.

Figure E-1 below identifies the activities that compose each of the eight phases. Each activity includes a description, as well as possible tailoring guidance, notes, and assumptions. Each activity is decomposed into tasks. Each task includes a description, information required to perform the task, the source of that information, and the information produced by the task [see Reference (k) for more information about tasks]. The VCM assumes that once a product is created, it is available for all subsequent activities; even though the product might not be identified as information that is required to perform the task.

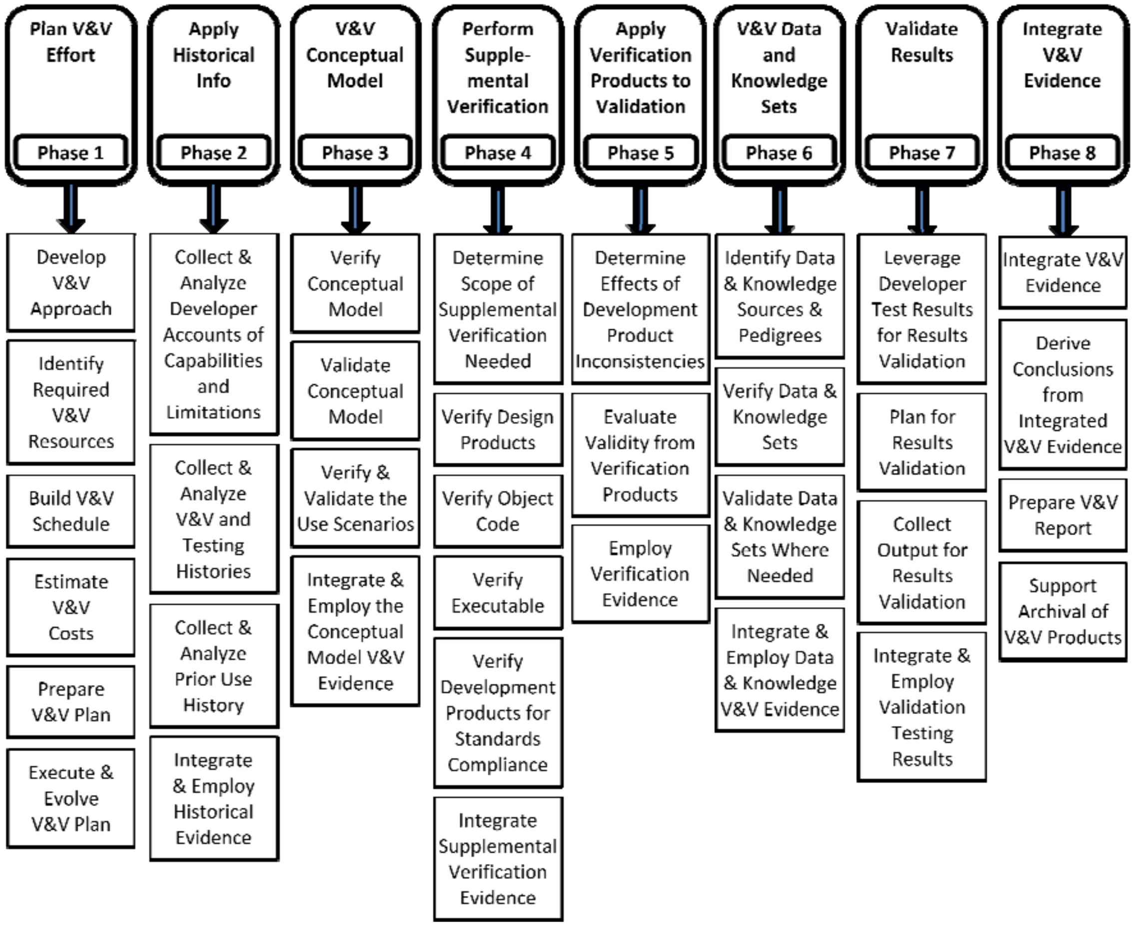


Figure E-1: Activities Associated with each VCM Phase (Graphical View)

Although many of the activities in Figure E-1 might appear sequential, the intention is not to restrict the implementation to any specific approach (e.g., waterfall, spiral, evolutionary, etc.). Rather, the diagram is meant to highlight the general activities from which V&V processes can be tailored to meet the needs of the M&S application based upon the intended use.

Table E-1 provides a tabular view of the information captured in Figure E-1.

Table E-1: Activities Associated with each VCM Phase (Tabular View)

#	TITLE
PHASE 1	**PLAN THE V&V EFFORT**
Activity 1.1	Develop the V&V Approach
Activity 1.2	Identify the Required V&V Resources
Activity 1.3	Build the V&V Schedule
Activity 1.4	Estimate the V&V Costs
Activity 1.5	Prepare the V&V Plan
Activity 1.6	Execute & Evolve the V&V Plan
PHASE 2	**APPLY RELEVANT HISTORICAL INFORMATION**
Activity 2.1	Collect & Analyze Developer Accounts of Capabilities & Limitations
Activity 2.2	Collect & Analyze the V&V and Testing Histories
Activity 2.3	Collect & Analyze the Prior Use History
Activity 2.4	Integrate & Employ the Historical Evidence
PHASE 3	**VERIFY & VALIDATE THE CONCEPTUAL MODEL**
Activity 3.1	Verify the Conceptual Model
Activity 3.2	Validate the Conceptual Model
Activity 3.3	Verify & Validate the Use Scenarios
Activity 3.4	Integrate & Employ the Conceptual Model V&V Evidence
PHASE 4	**PERFORM SUPPLEMENTAL VERIFICATION**
Activity 4.1	Determine the Scope of Supplemental Verification Needed
Activity 4.2	Verify the Design Products
Activity 4.3	Verify the Object Code
Activity 4.4	Verify the Executable
Activity 4.5	Verify the Development Products for Standards Compliance
Activity 4.6	Integrate the Supplemental Verification Evidence
PHASE 5	**APPLY THE VERIFICATION PRODUCTS TO VALIDATION**
Activity 5.1	Determine the Effects of Development Product Inconsistencies
Activity 5.2	Evaluate Validity from the Verification Products
Activity 5.3	Employ the Verification Evidence
PHASE 6	**VERIFY & VALIDATE THE DATA & KNOWLEDGE SETS**
Activity 6.1	Identify the Data & Knowledge Sources & Their Pedigrees
Activity 6.2	Verify the Data & Knowledge Sets
Activity 6.3	Validate the Data & Knowledge Sets Where Needed
Activity 6.4	Integrate & Employ the Data & Knowledge V&V Evidence

Table E-1: Activities Associated with each VCM Phase
(Tabular View) (continued)

#	TITLE
Phase 7	**VALIDATE THE RESULTS**
Activity 7.1	Leverage the Developer Test Results for Results Validation
Activity 7.2	Plan for Results Validation
Activity 7.3	Collect Output for Results Validation
Activity 7.4	Integrate & Employ the Validation Testing Results
PHASE 8	**INTEGRATE THE V&V EVIDENCE**
Activity 8.1	Integrate the V&V Evidence
Activity 8.2	Derive the Conclusions from the Integrated V&V Evidence
Activity 8.3	Prepare the V&V Report
Activity 8.4	Support Any Archival of the V&V Products

APPENDIX F: MATHEMATICAL LOGIC

This appendix contains five sub-appendices that discuss in more detail the techniques used in MURM for establishing the mathematical logic equation and evaluating weightings and probabilities for the various state tables of the factors.

Appendix F provides the main discussion of the mathematical logic that forms the MURM foundation. Appendices F1 — F5 provide additional detailed explanations of the mathematical logic behind the MURM concepts.

OVERVIEW

For MURM, the following definition is used for M&S Use Risk:

> *The probability that inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker.*

This definition is compatible with other definitions found in the literature, as indicated in a recent review by Hubbard[4] [Reference (l)]. This statement implies an unstated additional condition that is important for an unambiguous definition of use risk and should be addressed explicitly: both the inappropriate application and the unacceptable consequences occur (i.e., are true). A mathematical definition, corresponding to the definition above is as follows:

$$\text{M\&S Use Risk} = \text{UR} = p[\ (\text{Causes} \wedge \text{Effects}\ ) \wedge (\text{Causes} \Rightarrow \text{Effects})\ ]$$

$$= p(\text{Causes}) \times p(\text{Effects}) \times [\ 1 - p(\text{Causes}) + p(\text{Causes}) \times p(\text{Effects})\ ]$$

Where,

Causes ≡ the union of all the inappropriate applications of M&S Results for the intended use is true,

Effects ≡ the unacceptable consequences to the decision-maker are true

The symbol "$\wedge$" is the conjunction (i.e., "and") and the symbol "$\Rightarrow$" represents the implication (i.e., "if, then"). The probability, p(x), is the value of "x being true". This mathematical logic statement for M&S Use Risk provides a cogent and unambiguous definition when taken for exactly what it states (i.e., both "inappropriate application" and "unacceptable consequences" occur, and the latter is directly related to the former).

Therefore, M&S Use Risk is the probability that the inappropriate applications are true, that the unacceptable consequences are true, and that the unacceptable consequences are a result of the inappropriate applications.

DETERMINING p(CAUSES)

[4] In this work, Hubbard reviewed a century of risk terminology.

The probability of the causes of inappropriate application is discussed in detail because its value can be used to formulate and assess V&V Plans. A full discussion of the evaluation of Use Risk (UR), which includes not only p(Causes), but also p(Effects), will be addressed last.

MURM is applied individually to each M&S capability[5] that is required for the intended use. In the current analysis, three factors are defined as the causes of inappropriate applications: the clarity (or its lack) of the intended use, the M&S capability's importance to the decision, and the correctness of the V&V recommendation to employ or to not employ that capability of the M&S. The union of these factors leads to the following expression for determining the probability of the causes:

$$p(\text{Causes}) = p(C_1 \cup C_2 \cup C_3)$$

Where,

$C_1 \equiv$ Lack of clarity of intended use leading to misuse,

$C_2 \equiv$ Adverse impact on decision if capability is not achieved, and

$C_3 \equiv$ Incorrect recommendation to employ or not to employ M&S Results relative to that capability.

Factor C_1 (Clarity) is based on the understanding of the M&S intended use in a decision, and Factor C_2 (Importance) is based on the importance of the capability working properly and its impact on the decision. Factor C_3 (Confidence) is based on the recommendation arising from the examination techniques contained in a V&V Plan. The probability that the union of the factors is true is a function of the probability of each factor being true; the probability of a union is well defined[6] [Reference (m)], and for this case is:

$$p(C_1 \cup C_2 \cup C_3) = p(C_1) + p(C_2) + p(C_3) - p(C_1 \cap C_2) - p(C_1 \cap C_3) - p(C_2 \cap C_3) + p(C_1 \cap C_2 \cap C_3)$$

---

[5] Capability in MURM can be either individual parameters such as "detection capability of a particular kind of sensor, such as a sonar or radar," or the capability of a collection of related capabilities. MURM works both for the single-element capability and the combined aggregation of single-element capabilities into an aggregated capability.

[6] The basic equation of the probability of a union can be extended to any number of sets by recursion; the equation illustrated is for the union of three sets (C_1, C_2, and C_3), which, in the current discussion, represent the "cause" factors in the use risk construct.

where ∪ stands for the union, and ∩ for the intersection. In this equation, there are seven probabilities to determine: $p(C_1)$, $p(C_2)$, $p(C_3)$, $p(C_1 \cap C_2)$, $p(C_1 \cap C_3)$, $p(C_2 \cap C_3)$, and $p(C_1 \cap C_2 \cap C_3)$. Values for $p(C_2)$, $p(C)$, and $p(C_3)$ will be determined by using factor state tables. Simplifying assumptions are required for the other five. For example $p(C_2 \cap C_3) = p(C_2 | C_3) \times p(C_3)$ and is taken as $p(C_2) \times p(C_3)$ if the conditional is independent; similarly, $p(C_1 \cap C_2 \cap C_3) = p(C_1) \times p(C_2) \times p(C_3)$.

An application of MURM produces a summary product (a "scorecard") that brings together all the results. For a practitioner, it may the only product used. However, the detailed methods described in this section show the rigor in which the scorecard elements are developed so that in any tool employing this methodology, the practitioner can have confidence that the scorecard is based on well-founded analysis.

The scorecard contains the vector of capabilities (needed for the intended use), their corresponding clarity and/or importance, the associated recommendation attached to each capability (after the execution of the V&V Plan), the levels-of-confidence in the recommendations, and an overall projection of the use risk for each capability. At any stage of an analysis, some ingredients to compute value for a use risk will be present. An example of an abbreviated scorecard is illustrated in Table F-1.

Table F-1: Abbreviated Version of a MURM Scorecard

Capability	Importance	Activities & Examination Technique	Recommendation	Confidence (Percentile)
CAP_1	Medium	L: 3,3,1	TBD	12 PERCENTILE; VERY LOW
CAP_{2A}	High	L: 3,3,3	TBD	46 PERCENTILE; MEDIUM
CAP_{2B}	Medium	F: 3.3.5	TBD	29 PERCENTILE; LOW
CAP_3	Low	L: 3,3,5	TBD	79 PERCENTILE; HIGH
CAP_{N-1}				
CAP_N				

The mathematical definition of M&S Use Risk is based on several factors: information about the clarity of the intended use (Factor C_1), information about the importance of capability (Factor C_2), information coming from the V&V (Factor C_3), and the potential for unacceptable consequences to the decision-maker. The first three factors are the causes in M&S Use Risk, and the last factor is the effects. State tables for each factor are individually developed, and the probabilities associated with each level within a given state table are derived based on the maximum information entropy principle[7] [References (n) – (q)]. First, the details of building the cause factors are discussed because the MURM Scorecard and the V&V Plan can be analyzed using p(Causes). The full determination of use risk is completed by establishing a table for p(Effects) and computing M&S Use Risk as the probability of the mathematical logic expression of its definition.

STATE TABLE FOR FACTOR C_1: CLARITY OF INTENDED USE

The lack of clarity of the intended use in which an M&S capability is to be applied could lead to errors. For MURM, a three-level table[8] is shown in Table F-2; the requirements are that each level is mutually exclusive and the levels can be arranged in an agreed order-of-preference. During any period in a verification, validation, or accreditation process, updated information with respect to the use of an M&S capability would affect the state of Factor C_1 (Clarity), and subsequently, the M&S Use Risk calculation can be revised as more information is obtained.

Table F-2: State Table for Factor C_1 (Clarity), Assignment of $p(C_1)$

Factor Level[9]	Clarity of Intended Use	Level Weighting	$p(C_1)$
A	Lucid	1	0.167
B	Partial clarity	3	0.5
C	Unclear	5	0.833

[7] The maximum information entropy principle allows for unbiased weightings for factors when no information exists other than ranking according to some criteria. This principle has been extensively used in statistical thermodynamics and other applications, proving very helpful for avoiding unintended bias. Fundamental description of this concept may be found in works by Jaynes who promulgated its application half a century ago. See References (n) though (q). Latter references provide historical material about its application.

[8] The methodology may be expanded to any number of levels if they are clearly defined. For example, additional factor element ranges, uncertainties, sources, or the definition of a referent can be addressed as atoms in logical sentences that describe a level; the combinations of these sentences may produce several additional levels.

[9] The factor levels in this and subsequent state tables are labeled A, B, C, D, etc., to as many levels as defined as a shorthand notation; this notation should not be confused with that naming the factors [viz., Factor C_1 (Clarity), Factor C_2 (Importance), Factor C_3 (Confidence)].

The factor levels are listed from the "best" state to the "worst" state with respect to clarity. The attributes of clarity may include a specification of the ranges, uncertainties in the parameters, or other domain-related information to describe how the capability is to be applied. A "lucid" clarity would contain all that is required to eliminate ambiguity in the capability's intended use. In the next level down, some ambiguity may be present, and in the bottom level no or scant information is known.

STATE TABLE FOR FACTOR C_2: IMPORTANCE OF THE CAPABILITY TO THE DECISION

The execution of a V&V Plan may be constrained by resource availability so that all required capabilities cannot be examined or examined in the depth required by a maturity level.[10] Therefore, means to decide an appropriate distribution of plan resources will be needed. By determining the importance of a capability (again, always with respect to the intended use), an additional parameter will be available to facilitate an optimization strategy or at least organize the work in a V&V Plan by capability priority. Several techniques can be employed to arrange capability importance. Some are informal and potentially unnecessarily biased, while others have more rigor and reduce potentially unintended bias.[11]

For C_2, a rigorous approach that includes consequences and mitigation levels is used.[12] The consequences of a capability's inadequacy on the intended use's decision and the degree to which the mitigation of that inadequacy is achieved are combined. Table F-3 offers the list, in preferential order, along with a weighting (derived by maximum information entropy considerations—see Appendix F2) for use in trade-offs. The way to interpret the descriptions in Table F-3 is that <u>if</u> the M&S capability under examination is ineffective in its intended use, <u>then</u> the impact on the intended use's decision will be grave, serious, minor, or negligible.

[10] In MURM, maturity level is used to describe the quality of information used in M&S assessment by the VV&A process. The maturity level is an indication of information quality that impacts the confidence that may be placed in correctness of the recommendation about suitability of M&S Results related to the capability assessed. The VPMM is used as a basis for determining the maturity level associated with different levels of information quality. See Appendix C for additional details about VPMM.

[11] The objective is to avoid bias that is not supported by data that can be independently examined.

[12] Alternative approaches for constructing a C_2 state table are discussed in Appendix F3.

Table F-3: State Table for Factor C_2 (Importance), Assignment of $p(C_2)$

Factor Level	Consequence / Mitigation	Level Weighting	$p(C_2)$
A	Negligible consequence / Mitigation not required	1	0.038
B	Negligible consequence / Mitigation complete	3	0.115
C	Negligible consequence / Mitigation partial or Minor consequence / Mitigation complete	6	0.231
D	Negligible consequence / Mitigation impossible or Minor consequence / Mitigation partial or Serious consequence / Mitigation complete	11	0.423
E	Minor consequence / Mitigation impossible or Serious consequence / Mitigation partial or Grave consequence / Mitigation complete	17	0.654
F	Serious consequence / Mitigation impossible or Grave consequence / Mitigation partial	22	0.846
G	Grave consequence / Mitigation impossible	25	0.962

The assigned probability at each level must be between 0 and 1 and increases for the more troublesome levels; the ratios of these probabilities are specified by the level weighting (determined by the maximum information entropy principle), the results ranging from 0.038 through 0.962 in the order shown, in preferential order. For example, a Minor consequence/Mitigation partial combination is assigned a 0.423 probability of having an adverse effect if the Minor consequence/Mitigation partial is realized.[13] The way to interpret the probabilities in this table is that <u>if</u> an M&S capability is in the consequence/mitigation pairing state, <u>then the probability of adverse consequences</u> is as assigned; this <u>does not</u> mean the adverse consequences will definitely occur, but has that value of probability of occurring.

[13] As before, this is a result of the unbiased estimate using no other data than the order of the levels; the seven levels have a non-uniform distribution of possible combinations, thus the probability distribution sequence is non-uniform.

The grouping of several consequence/mitigation pairings is a result of the function chosen for the integration of these two sub-factors, as explained in Appendix F3.

STATE TABLE FOR FACTOR C_3: CONFIDENCE IN THE RECOMMENDATION ("TO USE" OR "NOT TO USE")

The V&V Plan: Activities, Tasks, and Techniques

The set of possible VV&A activities for examining an M&S is derived from the VCM. The VCM's activities were analyzed and organized into reasonable independent groupings. Many of the separate tasks in the VCM's original formulation can be grouped; in nearly all of those cases where options exist for individual task selection, the choices are driven by the availability of information or prior specification of the scope of the V&V effort and not by the goal of improving the certainty or coverage of the V&V evidence. The subactivities (i.e., the collections of tasks) were analyzed to determine their reasonable groupings and selections. Again, any valid selection options are driven by the availability of information or the goals of the V&V effort. Finally, grouping options at the activity level (i.e., groups of subactivities) lead to a tractable organization of major activities around which a recommendation confidence factor can be built:

1. Plan the V&V effort
2. Analyze historical information
3. Verify and validate the conceptual model
4. Perform supplemental verification
5. Use the verification products
6. Verify and validate the data and knowledge
7. Validate the M&S Results
8. Integrate the V&V evidence

It is assumed that all V&V efforts would begin with planning (the activity that produces the V&V Plan) and end with integrating the evidence (the activity that produces the V&V report). These required activities sandwich the activities that actually collect the V&V evidence (i.e., the activities numbered (2)-(7) above). It is also assumed that some form of data V&V would be required unless no data are available. Furthermore, the two verification activities, number (4) and number (6) can be collapsed into a single activity [called Development Product (DP) verification] because one would seldom, if ever, be performed without the other.

This process leaves the following four tailoring choices: number (2) historical record analysis, number (3) Conceptual Model (CM), DP verification, the combination of numbers (4) and (5), and number (7) results validation. Of course, because order is important, there are 24 (i.e., 4 factorial, or 4x3x2x1) possible combinations of four elements. However, the inherent causality of the M&S development lifecycle makes some of these orderings infeasible and others impractical. The opportunity to benefit from prior information eliminates still other combinations. What remains are the following combinations of these four elements into activity sets, listed roughly in order of improved evidence, labeled A through M.[14]

A. CM V&V

B. DP verification

C. historical record analysis or CM V&V + DP verification

D. historical record analysis + CM V&V

E. historical record analysis + DP verification

F. historical record analysis + CM V&V + DP verification

G. results validation

H. CM V&V + results validation

I. historical record analysis + results validation or DP verification + results validation

J. historical record analysis + CM V&V + results validation

K. historical record analysis + DP verification + results validation

L. CM V&V + DP verification + results validation

M. historical record analysis + CM V&V + DP verification + results validation

Note that two groups, items C and I, each contain two combinations because they are judged to represent equal value of evidence. Each individual activity provides evidence in the preferential order shown (i.e., CM V&V < DP verification < historical record analysis < results validation).

[14] These labels will correspond to the levels of the state table built for the Activities Selection sub-factor.

The foregoing construct of evidence value is translated into a mathematical logic that will facilitate the construction of factor tables and the weighted values of the states therein. First, the elemental tasks are defined as follows:

w = Conceptual Model Verification & Validation (CM V&V)

x = Development Product Verification (DP verification)

y = Historical Record Analysis

z = Results Validation

Using the symbols of the elemental tasks, the list of activity sets (A through M) are expressed in logical sentences as shown in Table F-4. This is designated as Factor (a), the table for activities selection. In Table F-4, positive atoms are bolded.

Table F-4: Factor(a) Activities Selection

Activity Set	Logical Sentence #1	Logical Sentence #2	Weighting	Adjusted Weighting
A	**w**^~x^~y^~z	-	1	100
B	~w^**x**^~y^~z	-	3	300
C	~w^~x^**y**^~z	**w**^**x**^~y^~z	6	600
D	**w**^~x^**y**^~z	-	9	900
E	~w^**x**^**y**^~z	-	11	1100
F	**w**^**x**^**y**^~z	-	13	1300
G	~w^~x^~y^**z**	-	15	1500
H	**w**^~x^~y^**z**	-	17	1700
I	~w^~x^**y**^**z**	~w^**x**^~y^**z**	20	2000
J	**w**^~x^**y**^**z**	-	23	2300
K	~w^**x**^**y**^**z**	-	25	2500
L	**w**^**x**^~y^**z**	-	27	2700
M	**w**^**x**^**y**^**z**	-	29	2900

The weighting of each state (i.e., row), shown in the fourth column, is determined by the methods of Appendix F2. The adjusted weighting will be explained once the second factor, Factor (b), is established in the next discussion. Note that the sentences expressing each activity set, the second and third columns, are completed by using the inverse expressions (i.e., ~x) of the missing tasks; these are provided for completeness and all completed sentences are mutually exclusive, which justifies the value of the weightings. These missing tasks will not influence the eventual computation of a weighted raw score (WRS) of an examination combination. The sets in levels C and I contain two sentences each, which are connected by an exclusive "or" — that is one expression or the other applies, but not both.

Factor (b), as expressed in Table F-5, provides for the level-of-technique selection for each of the elemental tasks, w, x, y, and z. The level of an examination technique is suggested by the VPPM level (see Appendix C). A planner is free to specify any of the Factor (b) five levels, A through E, which subsequently will be used in the computation of the WRS; however, good practice may restrict these selections (see Appendix E).

Table F-5: Factor(b) Technique Selection

Factor Level	Technique	Weighting	w Adjusted Weighting	x Adjusted Weighting	y Adjusted Weighting	z Adjusted Weighting
A	Use a LEVEL "1" technique	1	199	597	1194	2985
B	Use a LEVEL "2" technique	3	597	1791	3582	8955
C	Use a LEVEL "3" technique	5	995	2985	5970	14925
D	Use a LEVEL "4" technique	7	1393	4179	8358	20895
E	Use a LEVEL "5" technique	9	1791	5373	10746	26865

Factor (b) really represents four tables, one each for task elements w, x, y, and z. The weightings under the last four columns maintain the original relative weighting of the Factor (b) levels, but they are adjusted by multiplication by 199 x (1, 3, 6, and 15), respectively. This ensures that Factor (a) and Factor (b) have equal influence but also provides for the differential importance of elements w, x, y, and z.[15] The relative importance of w, x, y, and z are taken from Table F-4, where each singly appears in levels A, B, C, and G of that table.

The selection of an activity set for Factor (a) (i.e., A through M) and Factor (b) examination techniques (A through E), one for each elemental task that is present in the Factor (a) level leads to the formation of a WRS. This is the equation for the WRS:

WRS = [Factor (a) ADJ Weighting] + [Factor (b) ADJ Weighting]

where:

[Factor (b) ADJ Weighting] =

If(w = 'True', 1,0) * [w ADJ Weighting]

+ If(x = 'True', 1,0) * [x ADJ Weighting]

+ If(y = 'True', 1,0) * [y ADJ Weighting]

+ If(z = 'True', 1,0) * [z ADJ Weighting]

The score depends on Factor (a) and Factor (b) in the manner shown; depending of the level of Factor (a), only the weighting of the existing elemental tasks in that level is included. Based on the tables presented for Factor (a) and Factor (b), there are 1,295 combinations that can be formed through the various weightings. Not all values of WRS are unique—there are 484 unique values. Figure F-1 illustrates the distribution of the weighted raw scores.

[15] The adjusted weights for Factor (a) and Factor (b) arise because of the different number of levels in each factor. As a baseline, each factor should contain an equal influence to a weighted raw score, which is formed by summing weights taken from each table. The sum of the weightings (Table F-2, column 4) in Factor (a) is 199; that of Factor (b) (Table F-3, column 3) is 25. Because Factor (a) could be combined with each of the four elements, the adjusted values are formed by multiplying the weighting by 25x4 (=100); the weightings in Factor (b) are multiplied by 199 for the baseline and further multiplied by the relative weightings of w, x, y, and z (viz., 1,3,6, and 15).

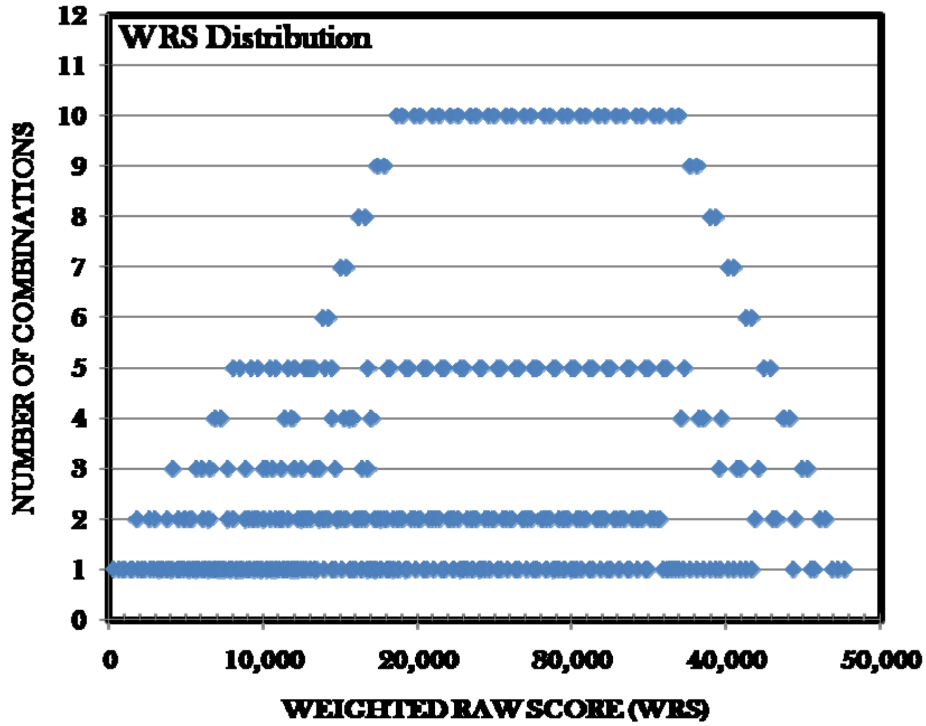


Figure F-1: Distribution of Weighted Raw Score

From this distribution, a WRS cumulative distribution, in percentile form, is derived. This cumulative distribution is employed to determine the confidence in a capability's recommendation. A higher percentile indicates a wider breath of evidence and a more focused depth of examination, thereby implying a high confidence in the resulting recommendation. Figure F-2 illustrates this cumulative distribution, and it is annotated in each quintile by a suggested phase to be associated with the confidence in the recommendation.

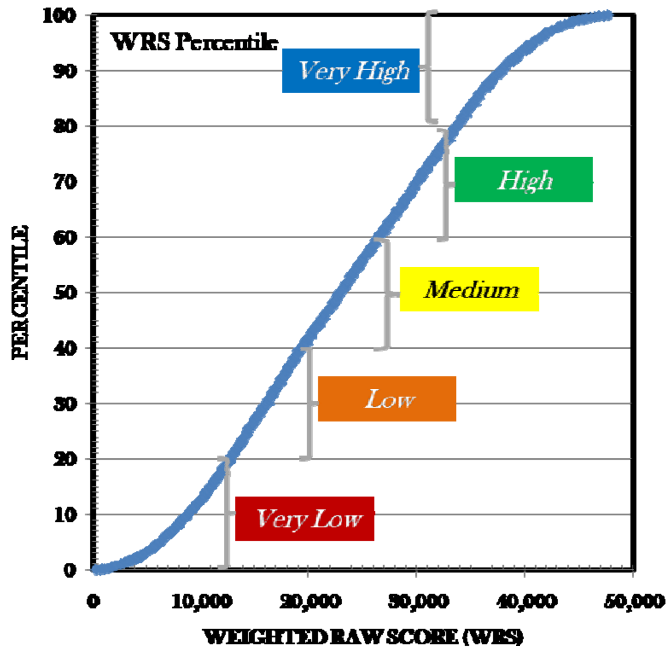


Figure F-2: Weighted Raw Score Percentiles based on Factor (a) and Factor (b)

Establishing the Factor C_3 (Confidence) State Table

If, as suggested that the result for recommendation confidence be reported in quintiles (i.e., percentiles 0-20, 20-40, etc.), then the recommendation confidence intervals will be grouped into five separate levels. A state table leading to an assignment of $p(C_3)$, Table F-6, may be built around the results of Figure F-2 (and the maximum information entropy arguments of Appendix F2).

Table F-6: State Table for Factor C_3 (Confidence), Assignment of $p(C_3)$

Factor Level	Recommendation Confidence	Level Weighting	$p(C_3)$
A	Confidence percentiles 80 to ≤100; Very high	1	0.05
B	Confidence percentiles 60 to <80; High	3	0.15
C	Confidence percentiles 40 to <60; Medium	5	0.25
D	Confidence percentiles 20 to <40; Low	7	0.35
E	Confidence percentiles 0 to <20; Very low	9	0.45

The relative weightings of each quintile are determined by the maximum information entropy principle knowing only the order of the levels. The $p(C_3)$ is determined from the baseline null "state of ignorance" of $p(C_3) = 0.5$, which is based on no information other than the recommendation can be either correct or incorrect.[16] The probability, $p(C_3)$, assigned to each level is interpreted as "the probability of being incorrect" with respect to a recommendation. As the factor levels add additional information from level E-to-level A, $p(C_3)$ decreases; the ratios of the $p(C_3)$s are preserved by the maximum information entropy principle. Because each probability must be less than 0.5 (the null state probability), and their ratios specified by the level weighting, the results are 0.05 through 0.45 in the preferential order, "best-to-worst." That is, a result with a very high confidence, a 0.05 probability of being incorrect, is derived. This is a result of the unbiased estimate using no other data than the order of the levels. In summary, as more Factor C_3 (Confidence) information is obtained by increased breath-of-evidence and more sophisticated examination, the level-of-confidence in making a correct recommendation increases and the probability of giving an incorrect recommendation decreases.

COMBINING FACTORS C_1, C_2, AND C_3 INTO P(CAUSES)

Table F-7 summarizes the derived probabilities for the three causes factors contained in Tables F-2, F-3, and F-6.

Table F-7: Summary of Derived Probabilities for p(Causes)

Factor Level	$p(C_1)$	$p(C_2)$	$p(C_3)$
A	0.167	0.038	0.05
B	0.5	0.115	0.15
C	0.833	0.231	0.25
D		0.423	0.35
E		0.654	0.45
F		0.846	
G		0.962	

[16] A probability of 0.5 is essentially a coin-flip.

The values for $p(C_1 \cup C_2 \cup C_3)$ are computed for all 105 combinations that can be formed by the values for the three factors show in Table F-7 using the following equation.[17]

$$p(C_1 \cup C_2 \cup C_3) = p(C_1) + p(C_2) + p(C_3) - p(C_1)\,p(C_2) - p(C_1)\,p(C_3) - p(C_2)\,p(C_3) + p(C_1)\,p(C_2)\,p(C_3)$$

The values for p(Causes) are dependent only on the order of the levels and their probabilities derived from their unbiased weightings from the three state tables (Tables F-2, F-3, and F-6). The only additional information that is used is the equation for $p(C_1 \cup C_2 \cup C_3)$.

Table F-8 contains several examples applying the foregoing method to assess an overall p(Causes). Several of the examples show trends on how each of the factors affects the overall level. Recall that Factor C_1 is the clarity of the intended use, Factor C_2 is the capability's importance to the decision, and Factor C_3 is the confidence in the recommendation ("to use" or "not to use").

[17] The equation assumes that C_1,C_2, and C_3 are uncorrelated so that $p(C_1 \cap C_2)$, $p(C_1 \cap C_3)$, $p(C_2 \cap C_3)$, and $p(C_1 \cap C_2 \cap C_3)$ can be calculated as $p(C_1)\,p(C_2)$, $p(C_1)\,p(C_3)$, $p(C_2)\,p(\,C_3)$, and $p(C_1)\,p(C_2)\,p\,C_3)$, respectively.

Table F-8: Examples of p(Causes) Calculation

Example	CLARITY C_1 Level [p(C_1)]	IMPORTANCE C_2 Level [p(C_2)]	CONFIDENCE C_3 Level [p(C_3)]	p(Causes) [p($C_1 \cup C_2 \cup C_3$)]
1	Lucid; A [0.167]	Grave/Partial; F [0.846]	Very high; A [0.05]	0.878
2	Lucid; A [0.167]	Grave/Partial; F [0.846]	Very low E [0.45]	0.930
3	Unclear; A [0.833]	Grave/Partial; F [0.846]	Very high; A [0.05]	0.976
4	Lucid; A [0.167]	Serious/Complete; D [0.423]	Very low E [0.45]	0.735
5	Unclear; A [0.833]	Serious/Complete; D [0.423]	Very low E [0.45]	0.947
6	Partial; B [0.5]	Serious/Complete; D [0.423]	Medium; C [0.25]	0.784
7	Lucid; A [0.167]	Serious/Complete; D [0.423]	Medium; C [0.25]	0.639
8	Lucid; A [0.167]	Serious/Complete; D [0.423]	Very high; A [0.05]	0.543

Several observations for these examples are as follows:

Example 1: Factors C_1 (Clarity) and C_3 (Confidence) are in very good shape. Factor C_2 (Importance) is the dominant factor. A major course-of-action for reducing use risk is to fully mitigate the M&S with respect to the subject capability, if possible, or downgrade the importance of the capability to the decision.

Example 2: Factor C_2 (Importance) is still dominant despite the worst condition of Factor C_3 (Confidence). Address factor C_2 (Importance) first if possible, and improve Factor C_3 (Confidence) next.

Example 3: Factor C_2 (Importance) is still dominant despite the better condition of Factor C_3 (Confidence). Address factor C_2 (Importance) first, if possible, and improve Factor C_1 (Clarity) next.

Example 4: The p(Causes) is Low. Factor C_3 (Confidence) is slightly more dominant than Factor C_2 (Importance). Improve Factor C_3 (Confidence), and then Factor C_2 (Importance), if possible.

Example 5: Factor C_1 (Clarity) is dominant and should be addressed first. Because the mitigation of the M&S is complete, the only way to reduce the influence of Factor C_2 (Importance) is to decrease the importance of the capability to the decision, if possible. Systematic improvement in factors C_1 (Clarity) and C_3 (Confidence), as shown in the next three examples (6, 7, and 8) can reduce p(Causes).

These tradeoffs are dealing with only one M&S capability. If a scorecard contains many M&S capabilities that are separately analyzed then more formal optimization techniques (e.g., linear programming, synthetic annealing) can be applied, especially if there are resource constraints. However, a "cost" function, which may include not only budget allocations, but also manpower, facilities, and specialized knowledge distributions, must be introduced into the analysis. The development of cost functions is not in the current scope of the MURM development.

USING THE SCORECARD TO TAILOR A V&V PLAN

With the computation of p(Causes), an expanded scorecard to track a V&V Plan can be constructed. Returning to the scorecard example introduced in Table F-1 and using the notations of the forgoing discussions for p(Causes), an expanded scorecard is illustrated in Table F-9.

Table F-9: Expanded Scorecard

Capability	Clarity	Importance	Activities & Examination Technique	Recommen-dation	V&V Evidence Confidence	M&S Use Risk
CAP_1	Lucid; A	Medium; D	SET L: 3,3,0,1	TBD	Very Low; E	P(Causes)=0.736 P(Effects)=TBD UR=TBD
CAP_{2A}	Partial; B	High; F	SET L: 3,3,0,3	TBD	Medium; C	P(Causes)=0.942 P(Effects)=TBD UR=TBD
CAP_{2B}	Lucid; A	Medium; D	SET F: 3,3,5,0	TBD	Low; D	P(Causes)=0.688 P(Effects)=TBD UR=TBD
CAP_3	Partial; B	Low; C	SET L: 3,3,0,5	TBD	High; B	P(Causes)=0.673 P(Effects)=TBD UR=TBD
…						
CAP_{N-1}						
CAP_N						

In the expanded scorecard, for each capability, the states of the cause factors are shown using the notations from Tables F-2 through F-8. The specific values for $p(C_1)$, $p(C_2)$, and $p(C_3)$ can be easily selected from Table F-7, and p(Causes) computed based on those values. The V&V Plan for each capability is noted in the "Activities & Examination Technique" column and uses a shorthand notation: the activity set (A through M) is noted and the levels of the examination technique for w, x., y, z immediately follows; if a task is not contained in that activity set then a "0" is shown in place of the examination level.

The scorecard would be continually updated as the V&V Planning and executions proceed, including the tradeoff of resources that would improve or change the various factor scores. At some point, the recommendations would be made (replacing the TBDs). In addition, a capability's effect on the decision may become known (i.e., p(Effects) would be set), and the computation for M&S Use Risk may be completed.

At times, it will be desirable to roll-up assessments such as shown in Table F-9 so a combination of capabilities is represented in an aggregate capability (the extreme example of such is the combination of all capabilities in an aggregate capability for the entire M&S). How such capabilities should be rolled-up will depend upon many considerations, and is likely to vary with circumstance. In some situations, the appropriate value of M&S Use Risk for an aggregate capability might be the maximum value for any of the capabilities in the combination of capabilities. In other situations, it might be some kind of average of the capabilities. In yet other situations, it might be determined in some other way.

COMPLETING THE DETERMINATION OF M&S USE RISK

The complete equation for M&S Use Risk is

$$UR = p(Causes) \times p(Effects) \times [1 - p(Causes) + p(Causes) \times p(Effects)]$$

The evaluation of p(Causes) is detailed above so that it can be used in V&V Planning. To complete the computation of M&S Use Risk, p(Effects), which is decision-maker dependent, must be specified. This is accomplished by building the associated state table for p(Effects), which is illustrated in Table F-10.

Table F-10: State Table for Effects Factor, Assignment of p(Effects)

Factor Level	Unacceptable Consequences to Decision-maker	Level Weighting	p(Effects)
A	Probability of unacceptable consequences is LOW	1	0.167
B	Probability of unacceptable consequences is MEDIUM	3	0.5
C	Probability of unacceptable consequences is HIGH	5	0.833

The details for the Effects factor may be intended-use-dependent and may vary from program to program. A general table is used here to illustrate the technique, which may be sufficient for most applications. However, more levels may be identified and the approach is flexible enough to accommodate expanded tables developed with the techniques of Appendix F2 or based on other data. Appendix F4 illustrates the derivation of an alternate p(Effects) table.

An illustration of the "lines of constant risk" for the full solution space of the M&S Use Risk equation is shown in Figure F-3. Each line represents a constant value of UR = p(Causes) x p(Effects) x [1 - p(Causes) + p(Causes) x p(Effects)] and divides the complex surface above the p(Causes)-p(Effects) plane into five equal areas.

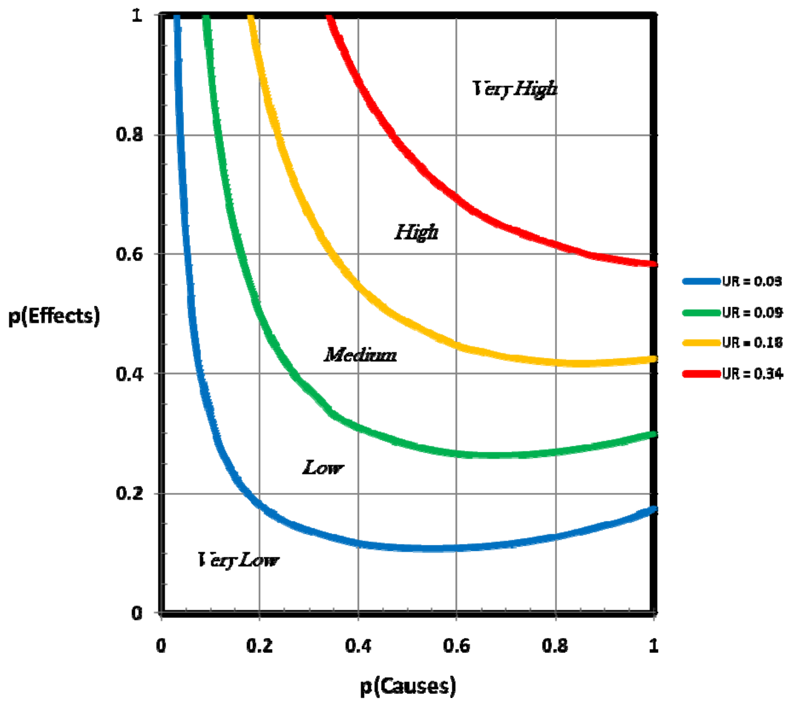


Figure F-3: M&S Use Risk (UR) Showing Lines of Constant Risk

The region in the lowest quintile (below the blue line to the axes) is labeled "Very Low" risk; the region between the blue-and-green lines is "Low" risk; those between the green-and-yellow lines, between the yellow and orange lines, and the remaining region above the orange line are "Medium," "High," and "Very High," respectively.

Using the information summarized in Table F-8 that contains eight examples, the computation for M&S Use Risk is completed for three of the five levels of p(Effects), as shown in Table F-11.

Table F-11: Computation of M&S Use Risk for Examples

Example	p(Causes) [p($C_1 \cup C_2 \cup C_3$)]	p(Effects)	UR
1	0.878	0.167	0.039; Low
		0.5	0.246; High
		0.833	0.624; Very high
2	0.930	0.167	0.035; Low
		0.5	0.249; High
		0.833	0.654; Very high
3	0.976	0.167	0.030; Low
		0.5	0.250; High
		0.833	0.680; Very high
4	0.735	0.167	0.048; Low
		0.5	0.232; High
		0.833	0.537; Very high
5	0.947	0.167	0.034; Low
		0.5	0.249; High
		0.833	0.664; Very high
6	0784	0.167	0.045; Low
		0.5	0.238; High
		0.833	0.568; Very high
7	0.639	0.167	0.050; Low
		0.5	0.217; High
		0.833	0.475; Very high
8	0.543	0.167	0.050; Low
		0.5	0.198; High
		0.833	0.411; Very high

SECTION SUMMARY

MURM was demonstrated in several steps. The methodology is performed on each individual M&S capability that is relevant to the intended use. First, the expression for M&S Use Risk was put into a mathematical logic form so that the probability of this statement could be evaluated using the probabilities of the elements, viz., the causes and the effects. Then a detailed analysis of three causes of M&S Use Risk were analyzed and state tables from which the probabilities of these causes could be assigned were derived. A scorecard is used to keep track of the relevant information that determines the probability of the causes, the probability of the effects, and the resulting M&S Use Risk. Tradeoff decisions for distributions of V&V resources or supporting decisions based on the intended use can be based on the MURM scorecard, which contains the calculated results of all these steps.

The mathematical details of this analysis are expected to be embedded in a tool so that a person applying this methodology will simply provide information in response to questions about the states of the causes and effects. The detailed calculations associated with the causes-effects relationship would be computed by the tool. This discussion has illustrated the rigorous mathematical basis that is contained in MURM to support its estimations of M&S Use Risk.

APPENDIX F1: DERIVATION OF THE MATHEMATICAL LOGIC EQUATION FOR M&S USE RISK

The definition of M&S Use Risk that has been adopted is as follows:

> *The probability that inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker.*

To put these words into a mathematical form that can be used in a numerical calculation of M&S Use Risk, we begin by parsing this statement:

> <u>The probability that</u> [(inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker) <u>AND</u> (that inappropriate application of M&S Results for the intended use occur) <u>AND</u> (unacceptable consequences to the decision-maker occur)].

In symbols (an explanation of the mathematical notation can be found in Appendix G):

Causes = C ≡ inappropriate application of M&S Results for the intended use

Effects = E ≡ unacceptable consequences to the decision-maker

$$\text{M\&S Use Risk} = p[(C \wedge E) \wedge (C \Rightarrow E)]$$

The value of M&S Use Risk, therefore, will be taken as the probability that the mathematical logic statement $[(C \wedge E) \wedge (C \Rightarrow E)]$ is true. In many "risk analyses" the multiplication of a cause ("failure") and effect ("impact") is taken as the value for risk. However, this approach is incomplete. Taken alone, the multiplication of two probabilities represents the probability of the coincidental occurrence of independent, non-mutually exclusive events. The intent of risk analyses is to evaluate a "cause-effect" relationship that is not produced by two independent events. Therefore, the complete definition requires the conjunction of the existence of causes and effects and the implication that those effects are produced by those causes as shown by the above equation.

To evaluate the probability of (C⇒E), the clause is put into disjointed canonical form. In disjointed canonical form, the implication is equivalent to the following:

$$C \Rightarrow E \equiv \sim C \vee E$$

where "~" is "not" and "∨" means "or".

This relationship is shown in the following Truth Table, where "T" is true and "F" is false:

C	E	~C	C ⇒ E	~C ∨ E
T	T	F	T	T
F	T	T	T	T
T	F	F	F	F
F	F	T	T	T

The complete construct for M&S Use Risk will give only one true for a risk definition while maintaining the condition that causes must be related to effects and not be independent:

C	E	C⇒E	C∧E	(C⇒E) ∧ (C∧E)
T	T	T	T	T
T	F	F	F	F
F	T	T	F	F
F	F	T	F	F

The requirement that C and E are true makes the construct true only. The probability for the full construct is computed by first obtaining the probability of C ⇒ E and multiplying it by the probability of C∧E:

$$p(C \Rightarrow E) \equiv p(\sim C \vee E) = p(\sim C) + p(E) - p(\sim C) \text{ x } p(E)$$

$$p(C \Rightarrow E) \equiv 1 - p(C) + p(E) - (1 - p(C)) \text{ x } p(E) = 1 - p(C) + p(E) - p(E) + p(C) \text{ x } p(E)$$

Simplifying: $p(C \Rightarrow E) \equiv 1 - p(C) + p(C) \text{ x } p(E)$

The probability of C∧E is p(C) x p(E), so the complete equation for M&S Use Risk is

$$\text{M\&S Use Risk} = p[(C \wedge E) \wedge (C \Rightarrow E)] = p(C) \times p(E) \times [\, 1 - p(C) + p(C) \times p(E) \,]$$

This construct will give only one true for a risk definition while maintaining the condition that causes must be related to effects and not be independent.

A plot for M&S Use Risk (UR) is illustrated in Figure F1-1, where the lines of constant M&S Use Risk (UR) are shown on the p(C)-p(E) plane. UR=1 occurs only at a single point (the upper right corner) and UR=0 along the abscissa and ordinate axes:

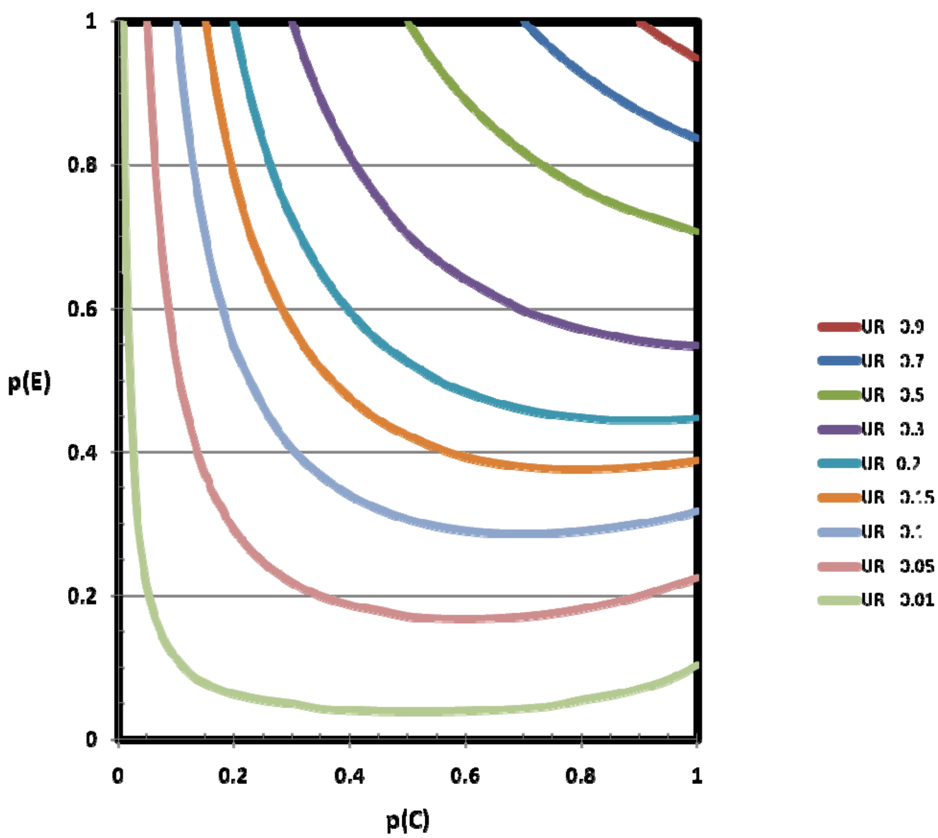


Figure F1-1: Lines of Constant M&S Use Risk (UR)

These lines of constant M&S Use Risk (UR) are projections of a three-dimensional surface onto a plane. The plot in Figure F1-2 represents the projection of this surface when it is divided into five equal areas. These quintiles are labeled into very low, low, medium, high, and very high risk respectively.

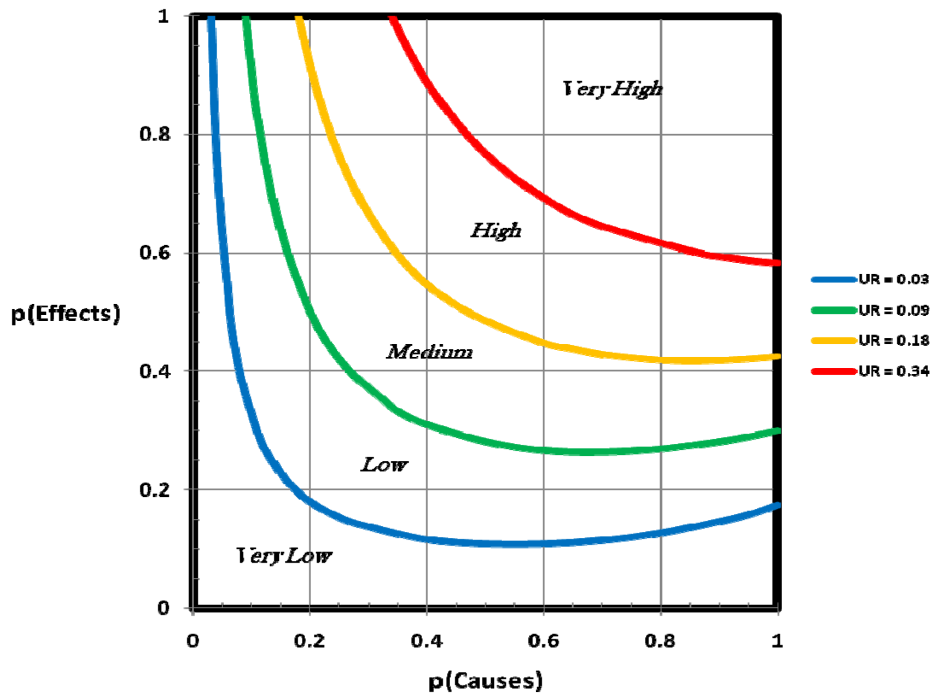


Figure F1-2: M&S Use Risk (UR) Showing Lines of Constant Risk

APPENDIX F2: USING MAXIMUM INFORMATION ENTROPY TO ESTABLISH WEIGHTINGS

The principle of maximum information entropy[18] provides an additional condition for determining the weightings of various levels in a state table. How this concept is employed to analyze and/or establish weightings is illustrated through several examples.

FORMING AND WEIGHTING ORDERED LISTS

Consider a system that has three classes (labeled 3rd Class, 2nd Class, and 1st Class), in increasing order of preference, as show in Figure F2-1. That is, members of the 1st Class are considered better ("more valuable") than those of the 2nd or 3rd Classes, and members of the 2nd Class are considered better than those of the 3rd Class.

Figure F2-1: Three Classes in Preferential Order

To assign relative weightings to the classes, knowing only the order of preference, line up the classes over an interval in their order of preference (Figure F2-2):

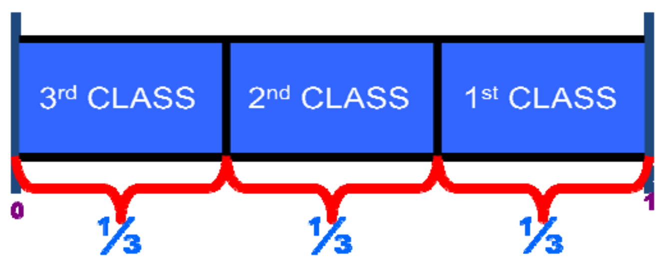


Figure F2-2: Classes Arranged in Order Over 0-To-1 Interval

By inspection, the interval between the (normalized) limits is broken into equal parts, and each class is the centroid value of their respective intervals. For equal partitions, these centroids are 1/6, 1/2, and 5/6 on the 0-to-1 interval, and

[18] See Jaynes, Reference (o)

have a ratio of 1:3:5, the weightings proposed to use for analyses, barring any detailed information (data) otherwise. These weightings give the most unbiased result based on the information available (i.e., only the order of preference) because they produce the maximum information entropy of the system.

The equation for information entropy is

$$S = \Sigma\, p_i \log_2 [1/p_i]$$

where S is information entropy, and p_i is the probability for the i^{th} class.

For the arrangement in Figure F2-2, which corresponds to the maximum information entropy, the value is

$$S_{max} = 1/3 \log_2 [3] + 1/3 \log_2 [3] + 1/3 \log_2 [3] = 1.5849$$

To illustrate the information entropy values of other partitions, consider two others (Figure F2-3 and Figure F2-4) that are not equal: For the Figure F2-3 partition, $S = 1/6 \log_2 [6] + 1/3 \log_2 [3] + 1/2 \log_2 [2] = 1.4591$, with centroids at 1/12, 1/3, and 3/4. The relative weightings are 1: 4: 9. Note that the value for the information entropy of this partition is less than the 1.5849 of the equal partitions case. A lower value for the information entropy implies that more is known about the system—in this case, some additional information would be needed to justify the 1:4:9 ratios or the size of the partitions.

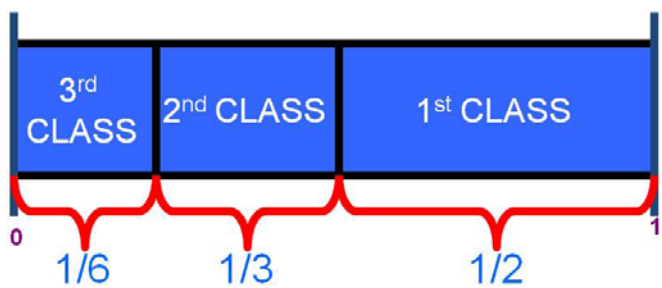


Figure F2-3: A Different Arrangement for the Three Classes

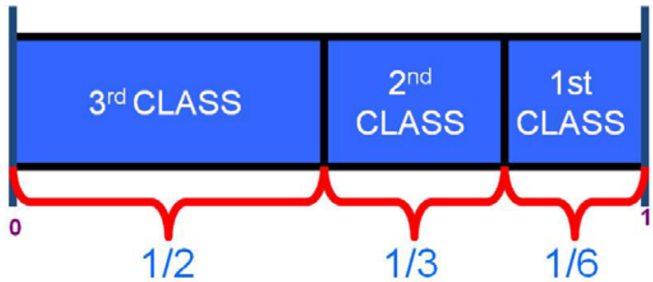


Figure F2-4: Another Partition

For the Figure F2-4 partition, S = 1/2 $\log_2$ [2] + 1/3 $\log_2$ [3] + 1/6 $\log_2$ [6] = 1.4591, with centroids at 1/4, 2/3, and 11/12. The relative weightings are 1: 2.667:3.667. This looks at the unequal partitions of the previous case, but with the size of the intervals reversed.

Both partitions have the same information entropy; however, the relative weightings of the classes are different because of the order in which the unequal intervals are distributed. In either case, the lower value of the information entropy implies that there is more information about the system than the equal partition case. In the first partition, being a member of a higher class is increasingly more valuable (the weighting going up as a square); in the second partition, there is a case of diminishing returns—the first step from 3rd to 2nd class gives a return of 2.667-to-1, but a second step from 2nd to 1st class is an incremental return of 1.375-to-1.

An analyst may be tempted to weigh three classes as 1:2:3. It is instructional to see what using class weightings in the ratio 1: 2: 3 implies. As it turns out, these ratios represent a limiting case of the illustrated partition (Figure F2-5), which has an entropy S = 1/2 $\log_2$ [2] + 1/200 $\log_2$ [200] + 0.495 $\log_2$ [2.02] = 1.0404, with centroids at 1/4, 0.5025, and 0.7525. The relative weightings are 1: 2.01 : 3.01.

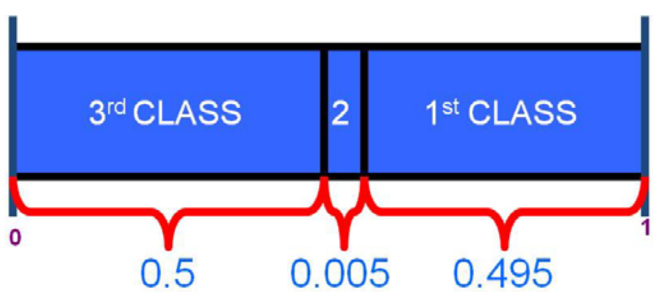


Figure F2-5: Partition for 1-2-3

As the 1st class interval is increased to approach 1/2, the 2nd class shrinks to a point at 0.5, and in the limit, the entropy is 1. Compare this limiting value to that of equal partitions and the two other distorted cases considered above. Using weightings in the ratio of 1:2:3 implies a knowledge about the system with three classes that is *very much more* than what is known in the equal partition case and in those cases that are not even close to equal partitions (i.e., an information entropy of 1 or 1.0404 is substantially less than 1.4591 or 1.5849). If, in an analysis, we do not possess information beyond that of the order of preference of the three classes, then the most unbiased ratio to choose is 1:3:5, the most justifiable weightings based on our current information.

To derive weightings from a compounding of individual elements into a smaller number of levels, the technique is to list the elements in order and divide them by the level partitions (see Figure F2-6). The relative weightings for all the elements in a partition are the ratios of the centroids in that division.

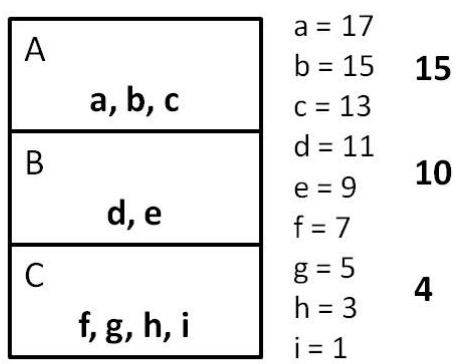


Figure F2-6: Weighting Compound Levels

The nine independent elements arranged a-through-i would have system information entropy of 3.1699 (i.e., $\log_2 9$). When the triage is performed, information is added to the system and the entropy decreases:

$p_A = 3/9$; $p_B = 2/9$; $p_C = 4/9$

$$S = (3/9)\log_2(9/3) + (2/9)\log_2(9/2) + (4/9)\log_2(9/4) = 1.5305$$

Throughout the state tables discussed in Appendix F, the principle of maximum information entropy is applied to determine appropriate weightings of factors based solely upon the order of preference of those factors, modified by any additional available information (such as the logical sentence structure within a level).

APPLICATION TO MURM

These methods are applied in building the state tables for the MURM factors. For MURM the lists proceed from best to worst (designated by levels A, B, C, etc.); the diagram of the layout of the more valuable classes (those associated with a lower probability of cause or effect) are on the left. Once the partitions are established, the centroids are taken as the derived probability associated with that state's level.

ANALYSIS FOR FACTOR C_1 (Clarity):

Table F2-1 lists the states for Factor C_1 (Clarity):

Table F2-1: States for Factor C_1 (Clarity)

Factor Level	Clarity of Intended Use
A	Lucid
B	Partial clarity
C	Unclear

The diagram for this arrangement is shown in Figure F2-7. The centroids of each level are shown and their ratios correspond to the relative weightings of the states. The associated probabilities for the factor being true maintain this ratio and are 0.16667, 0.5, and 0.83333, for levels A, B, and C, respectively.

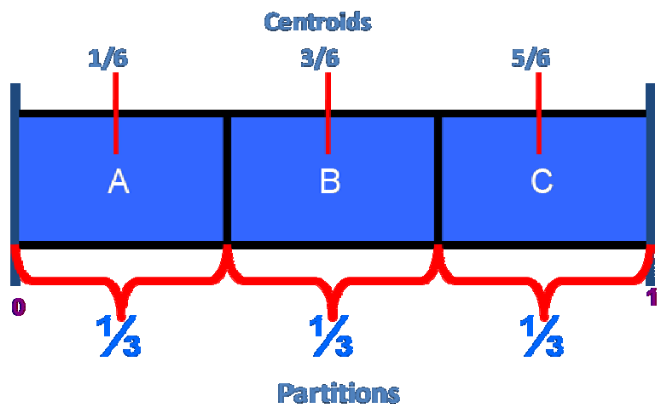


Figure F2-7: Diagram for the Factor C_1 (Clarity) State

ANALYSIS FOR FACTOR C_2 (Importance)

Table F2-2 lists the states for Factor C_2 (Importance):

Table F2-2: States for Factor C_2 (Importance)

Factor Level	Consequence / Mitigation
A	Negligible consequence / Mitigation not required
B	Negligible consequence / Mitigation complete
C	Negligible consequence / Mitigation partial OR Minor consequence / Mitigation complete
D	Negligible consequence / Mitigation impossible OR Minor consequence / Mitigation partial OR Serious consequence / Mitigation complete
E	Minor consequence / Mitigation impossible OR Serious consequence / Mitigation partial OR Grave consequence / Mitigation complete
F	Serious consequence / Mitigation impossible OR Grave consequence / Mitigation partial
G	Grave consequence / Mitigation impossible

The diagram for this arrangement is shown in Figure F2-8. Note that the partitions are not uniform because levels C and F contain two mutually exclusive possibilities that are equally weighted, and levels D and E contain three mutually exclusive possibilities that are equally weighted. The centroids of each level are shown and their ratios correspond to the relative weightings of the states. The associated probabilities for the factor being true maintain this ratio and are 0.038, 0.115, 0.231, 0.423, 0.654, 0.846, and 0.962 for levels A, B, C, D, E, F, and G, respectively.

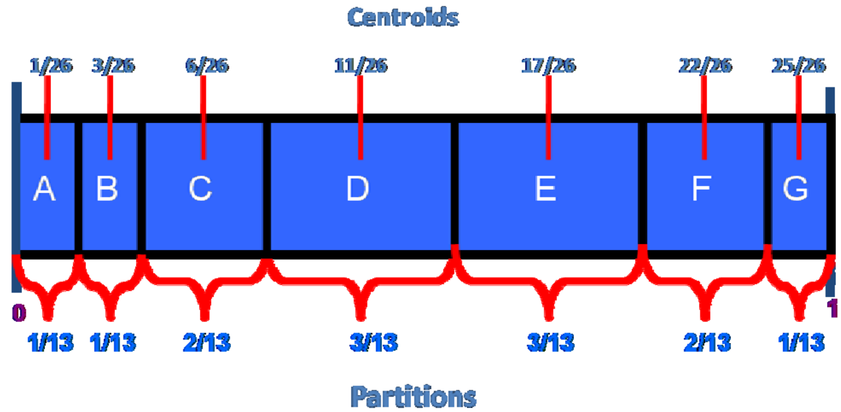


Figure F2-8: Diagram for the Factor C_2 (Importance) State

ANALYSIS FOR FACTOR C_3 (Confidence):

Table F2-3 lists the states for Factor C_3 (Confidence):

Table F2-3: States for Factor C_3 (Confidence)

Factor Level	Recommendation Confidence
A	Confidence percentiles 80 to ≤100; Very high
B	Confidence percentiles 60 to <80; High
C	Confidence percentiles 40 to <60; Medium
D	Confidence percentiles 20 to <40; Low
E	Confidence percentiles 0 to <20; Very low

The diagram for this arrangement is shown in Figure F2-9. The centroids of each level are shown and their ratios correspond to the relative weightings of the states. Because the decision associated with factor two is binary (i.e., correct or incorrect), the associated probabilities are 0.05, 0.15, 0.25, 0.35, and 0.45, for levels A, B, C, D, and E, respectively, and these probabilities maintain the ratios of the centroids, 1:3:5:7:9.

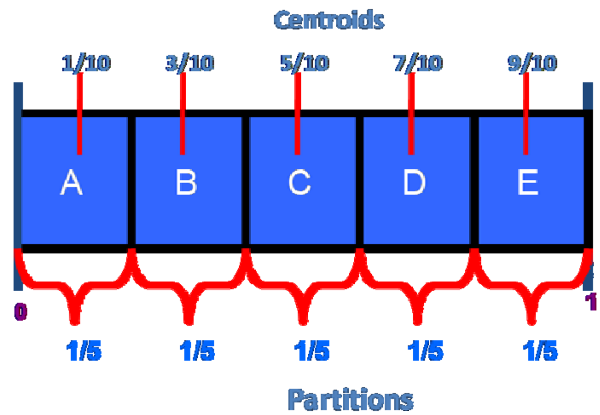


Figure F2-9: Diagram for the Factor C_3 (Confidence) State

ANALYSIS FOR FACTOR (a)

Table F2-4 lists the states for Factor (a), which is employed in the development of Factor C_3 (Confidence).

Table F2-4: States for Factor C_3 (Confidence) Factor (a)

Sentence	Logical Sentence #1	Logical Sentence #2
A	**w**^~x^~y^~z	-
B	~w^**x**^~y^~z	-
C	~w^~x^**y**^~z	**w**^**x**^~y^~z
D	**w**^~x^**y**^~z	-
E	~w^**x**^**y**^~z	-
F	**w**^**x**^**y**^~z	-
G	~w^~x^~y^**z**	-
H	**w**^~x^~y^**z**	-
I	~w^~x^**y**^**z**	~w^**x**^~y^**z**
J	**w**^~x^**y**^**z**	-
K	~w^**x**^**y**^**z**	-
L	**w**^**x**^~y^**z**	-
M	**w**^**x**^**y**^**z**	-

To develop the weightings for this factor, a diagram, Figure F2-10, similar to Figure F2-6, is constructed. Here the logical sentences are arranged from A to M; levels C and I each contain two mutually exclusive sentences, and the remaining levels contain one each. The weightings are shown in the rightmost column.

Sentence	# Logical Sentences	MIE Weighting	Median MIE Weighting
A	1	1	**1**
B	1	3	**3**
C	2	5,7	**6**
D	1	9	**9**
E	1	11	**11**
F	1	13	**13**
G	1	15	**15**
H	1	17	**17**
I	2	19,21	**20**
J	1	23	**23**
K	1	25	**25**
L	1	27	**27**
M	1	29	**29**

Figure F2-10: Weightings calculation for Factor (a)

ANALYSIS FOR FACTOR (b)

Table F2-5 lists the states for Factor (b), which is employed in the development of Factor C_3 (Confidence).

Table F2-5: States for Factor (b)

Maturity Level	Technique
A	Use a LEVEL "1" technique
B	Use a LEVEL "2" technique
C	Use a LEVEL "3" technique
D	Use a LEVEL "4" technique
E	Use a LEVEL "5" technique

To develop the weightings for this factor, a diagram, Figure F2-11, similar to Figure F2-6, is constructed. The result is straightforward because each level contains only one logical sentence. The weightings are shown in the last column.

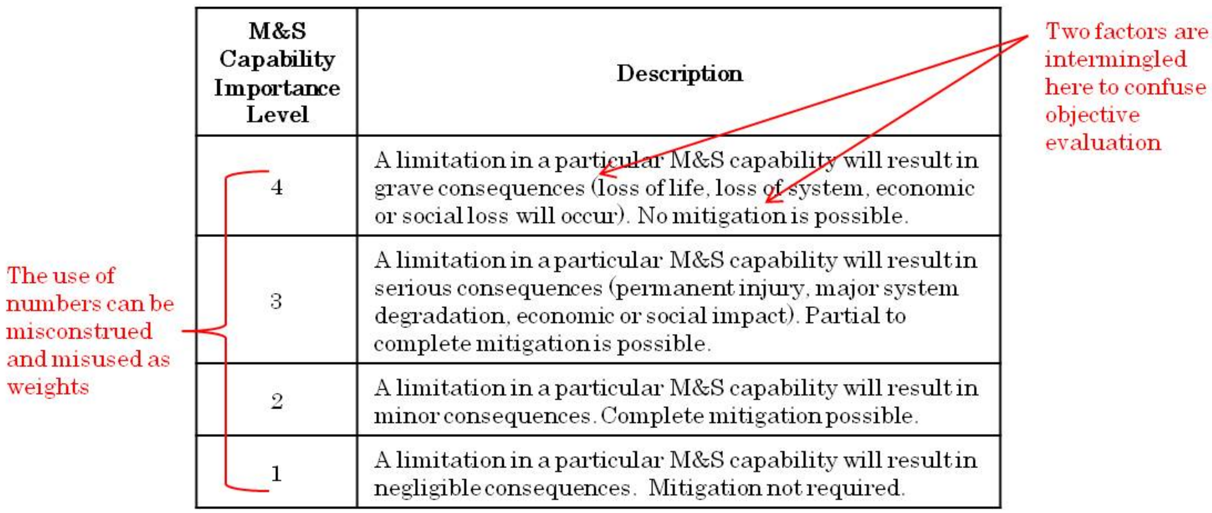
Sentence	# Logical Sentences	MIE Weighting	Median MIE Weighting
A	1	1	**1**
B	1	3	**3**
C	1	5	**5**
D	1	7	**7**
3	1	9	**9**

Figure F2-11: Weightings Calculation for Factor (b)

The analyses for weightings and probabilities in the remainder of this appendix are formed in similar ways to those demonstrated for Factors C_1 (Clarity), C_2 (Importance), and C_3 (Confidence).

APPENDIX F3: ANALYSIS FOR THE IMPORTANCE FACTOR STATE TABLE

A table attempting to specify the importance of an M&S capability appears on page 18 of reference (d) and was originally adapted from IEEE Std 1012™-2004 [Reference (r)]. Use of this table as it appears may lead to biased results in determining importance (see Figure F3-1). In the following discussion, the elements of the table are re-formulated and the analysis it contains is reformed by the Maximum Information Entropy (MIE) principle to yield a more universal way to assign M&S importance.



M&S Capability Importance Level	Description
4	A limitation in a particular M&S capability will result in grave consequences (loss of life, loss of system, economic or social loss will occur). No mitigation is possible.
3	A limitation in a particular M&S capability will result in serious consequences (permanent injury, major system degradation, economic or social impact). Partial to complete mitigation is possible.
2	A limitation in a particular M&S capability will result in minor consequences. Complete mitigation possible.
1	A limitation in a particular M&S capability will result in negligible consequences. Mitigation not required.

Figure F3-1: Determining M&S Importance

This table mixes two factors (consequences and mitigations) in such a way that it only addresses part of the solution space and assigns M&S capability importance levels that are biased without supporting data. The intention is to assign an importance to each capability based on the view of the impact on the decision if the recommendation turns out that the M&S capability is ineffective. Such an assignment will help identify highly critical capabilities for which examinations should be rigorous.

In Figure F3-2, the two factors are separated into tabular form, weighted each according to MIE principles (see Appendix F2), and combined by the sum of their weightings to derive a combined score (URS) for the state of M&S consequences (i.e., it being ineffective) and its mitigation. Here it is assumed that mitigation attempts to align the M&S to its intended use.

				Consequences (Columns)			
				A	B	C	D
			W(C)	1	3	5	7
		W(M)		Negligible	Minor	Serious	Grave
Mitigation (Rows)	A	1	Not Required	2			
	B	3	Complete	4	6	8	10
	C	5	Partial	6	8	10	12
	D	7	Impossible	8	10	12	14

Legend: W(C) = consequence weight; W(M) = mitigation weight

Figure F3-2: Separating the Two Factors Into the Full Space

It is noted that mitigation would <u>not be required</u> for negligible consequences (however, in the state space, mitigations could be applied). Other consequences would have some form of mitigation ("complete", "partial", or "impossible") applied. The language for consequences and mitigations contained in the source table is employed. Conversely, for non-negligible consequences, some level of mitigation would be required ("not required" is not allowed).

The reformed analysis leads to the state table, Table F3-1, for M&S importance.

Table F3-1: A Complete Analysis for M&S Importance Indicating All Combinations

Importance Level	W(C) + W(M)	Column/ Row	Consequence/ Mitigation		Level Weighting	Percentile	p
A	2	AA	Negligible	Not Required	1	0.00	0.038
B	4	AB	Negligible	Complete	3	8.33	0.115
C	6	AC,BB	Negligible Minor	Partial Complete	6	25.00	0.231
D	8	AD,BC,CB	Negligible Minor Serious	Impossible Partial Complete	11	50.00	0.423
E	10	BD,CC,DB	Minor Serious Grave	Impossible Partial Complete	17	75.00	0.654
F	12	CD,DC	Serious Grave	Impossible Partial	22	91.67	0.846
G	14	DD	Grave	Impossible	25	100.00	0.962

The reason for adding the two factor relative weightings, W(C) + W(M), to form a URS is because both are used to assign a combination's importance and each factor is assumed to be independent. An alternate way would be to multiply the values of the two factors; this can be shown to imply a probability (of suitability failure) for the consequence/mitigation combination. The form of the function combining C & M needs to be checked out to ensure unbiased analysis.

By this derived table, an M&S with a limitation that is impossible to mitigate and that leads to grave consequences carries the highest importance. M&S with negligible consequences rank lower in importance (as expected), depending on the level of mitigation.

The weightings used in the above tables are derived directly from the MIE principle. Any other weightings would need to be justified by data (not opinion) that can be independently analyzed. The allowable combinations of consequences and mitigations can be adjusted and may lead to different weightings. For example, the "not required" mitigation level could be fused with the "impossible" level to make a combined level entitled "impossible or none." The three mitigation levels in order of best to worst would then be "complete," "partial," "impossible or none," and carry MIE weightings of 1, 3, and 5, respectively; the other numbers in the table would be accordingly adjusted and the relative importance weightings would change.

In Figure F3-3, the cumulative distribution (as a percentile) for M&S Importance is illustrated (note that it is a discrete distribution).

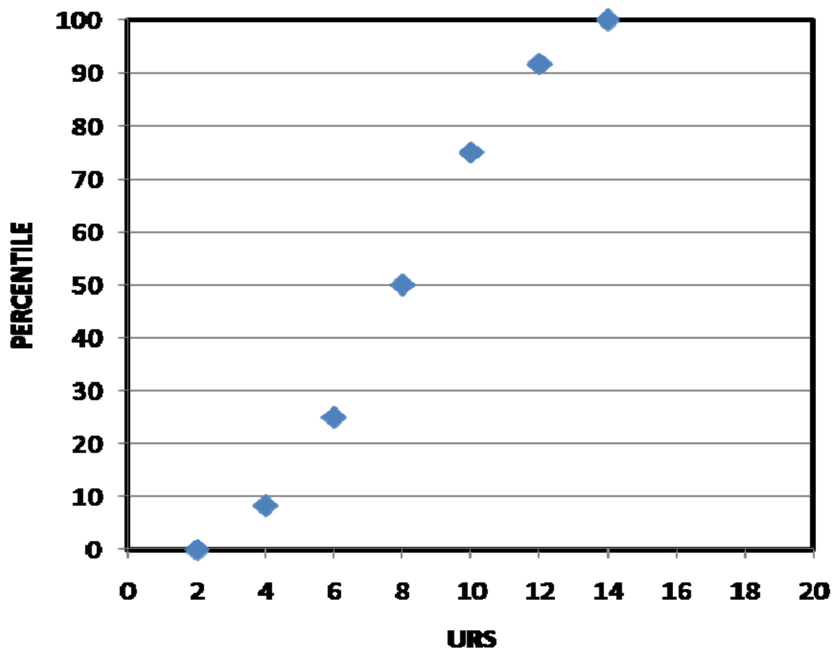


Figure F3-3: M&S Importance (Percentile as a Function of URS)

ALTERNATE FORMULATIONS

The preceding discussion for M&S Importance used a two sub-factor relationship that depended on the consequences and mitigation of the M&S capability. If this approach seems too difficult to implement with the available information, an alternative table can be based on any other arrangement of importance factors. Table F3-2 is an example of a simpler importance factor table. NOTE THAT ANY CHANGE in the importance factor table that affects the values for $p(C_2)$ (i.e., replacing Table F-4 in the main discussion) will require that the analysis contained in Tables F-7, F-8, F-9 and F-11 and Figure F-2 be revised to reflect the impact of the new values. The revisions are straightforward.

Table F3-2: IMPORTANCE Based on Consequence to the Intended Use Decision IF the Capability Needs a Limitation

Importance Level	Description	Weighting	$p(C_2)$
A	A limitation in this capability will result in NEGLIGIBLE consequences	1	0.125
B	A limitation in this capability will result in MINOR consequences	3	0.375
C	A limitation in this capability will result in SERIOUS consequences	5	0.625
D	A limitation in this capability will result in GRAVE consequences	7	0.875

Other similar tables may be constructed by separating the capabilities into various classes (by any means) and computing the relative weightings for each class by the method demonstrated in Appendix F2 (and shown here as Figure F3-4). The values for $p(C_2)$ for each importance level are derived as a function of the weightings.

Sentence	# Logical Sentences	MIE Weighting	Median MIE Weighting
A	4	1,3,5,7	**4**
B	2	9,11	**10**
C	3	13,15,17	**15**

Figure F3-4: Weighting Compound Levels

An importance factor table derived from Figure F3-4 is shown in Table F3-3. Here Class A represents the lowest importance group, and Class C represents the highest importance group.

Table F3-3: IMPORTANCE Based on Organization of Capabilities into Three Classes (Figure F3-4)

Importance Level	Description	Weighting	$p(C_2)$
A	Lowest members: f, g, h, and i	4	0.2
B	Medium members: d and e	10	0.5
C	Highest members: a, b, and c	15	0.75

Again, any changes in the values of $p(C_2)$ from those shown in Table F-5 of the main discussion will require changes to Tables F-7, F-8, F-9, and F-11 of the main discussion.

APPENDIX F4: ALTERNATE p(EFFECTS) FACTOR TABLE

For the purposes of illustrating the initial development of MURM, a placeholder table, as shown in Table F4-1, was used to compute p(Effects):

Table F4-1: State Table for Effects Factor, Assignment of p(Effects)

Factor Level	Unacceptable Consequences to Decision-maker	Level Weighting	p(Effects)
A	Probability of unacceptable consequences is LOW	1	0.167
B	Probability of unacceptable consequences is MEDIUM	3	0.5
C	Probability of unacceptable consequences is HIGH	5	0.833

As more information about the decision-maker's situation becomes known, other effects factor tables may be derived and used. Below is an illustration of the development of a more detailed table using information about M&S Impact and M&S Reliance. Two sub-factors are considered: (a) the intended use area and (b) the dependency (of the decision) on the M&S capability. Figure F4-1 illustrates the combination of these two factors into an unweighted raw score (URS = w(UA) + w(D), where UA is use area and D is dependency).

				M&S IMPACT (INTENDED USE)				
				A	B	C	D	E
				1	3	5	7	9
				Single Low Risk Area	Single Medium or Low Risk Area	Multiple Medium or Low Risk Areas	Single High Risk Area	Multiple High Risk Areas
M&S RELIANCE (DEPENDENCY)	A	1	Only	8	10	12	14	16
	B	3	Primary	6	8	10	12	14
	C	5	Secondary	4	6	8	10	12
	D	7	Supplemental	2	4	6	8	10

Figure F4-1: Construction of an Unweighted Raw Score for Combinations of Intended Use Area and Dependency on M&S

This procedure is similar to that shown in Appendix F3 for the importance factor based on consequence-mitigation values. The analysis leads to Table F4-2, which defines the Effects levels based on URS, their respective weightings, percentiles, and the derived p(Effects).

Table F4-2: Effects Factor Table Based On Intended Use Area and Dependency on M&S

EFFECTS Level	W(UA) + W(D)	Column/ Row	Dependency/Use Area		Level Weighting	Percen-tile	p
A	2	AA	Supplemental	Single LOW	0.00	1	0.025
B	4	AB,BA	Supplemental Secondary	Single MEDIUM Single LOW	10.53	4	0.100
C	6	AC,BB, CA	Supplemental Secondary Primary	Multiple MED/LOW Single MEDIUM Single LOW	26.32	9	0.225
D	8	AD,BC, CB,DA	Supplemental Secondary Primary Only	Single HIGH Multiple MED/LOW Single MEDIUM Single LOW	47.37	16	0.400
E	10	AE,BD, CC,DB	Supplemental Secondary Primary Only	Multiple HIGH Single HIGH Multiple MED/LOW Single MEDIUM	68.42	24	0.600
F	12	BE,CD, DC	Secondary Primary Only	Multiple HIGH Single HIGH Multiple MED/LOW	84.21	31	0.775
G	14	CE,DD	Primary Only	Multiple HIGH Single HIGH	94.74	36	0.900
H	16	DE	Only	Multiple HIGH	100.00	39	.975

For completeness, the percentile-URS distribution for this new effects table is shown in Figure F4-2.

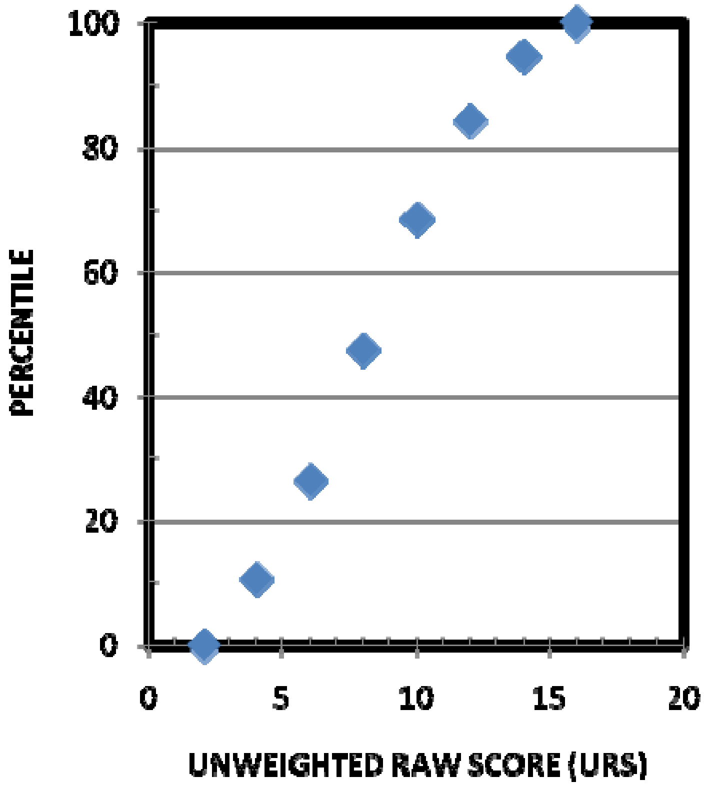


Figure F4-2: Percentile Distribution for Effects Factor

Table F4-3 is the effects factor table (corresponding to Table F-10 of main discussion) based on this new analysis:

Table F4-3: State Table for Effects Factor, Assignment of p(Effects)

Factor Level	Probability of Unacceptable Consequences to Decision-Maker Based on Dependency/Use Area	Level Weighting	p(Effects)
A	Supplemental Use/Single Low Risk Area	1	0.025
B	(Supplemental Use/Single Medium Risk Area) or (Secondary Use/Single Low Risk Area)	4	0.100
C	(Supplemental Use/Multiple Med-Low Risk Area) or (Secondary Use/Single Medium Risk Area) or (Primary Use/Single Low Risk Area)	9	0.225
D	(Supplemental Use/Single High Risk Area) or (Secondary Use/Multiple Med-Low Risk Area) or (Primary Use/Single Medium Risk Area) or (Only Use/Single Low Risk Area)	16	0.400
E	(Supplemental Use/Multiple High Risk Area) or (Secondary Use/Single High Risk Area) or (Primary Use/Multiple Med-Low Risk Area) or (Only Use/Single Medium Risk Area)	24	0.600
F	(Secondary Use/Multiple High Risk Area) or (Primary Use/Single High Risk Area) or (Only Use/Multiple Med-Low Risk Area)	31	0.775
G	(Primary Use/Multiple High Risk Area) or (Only Use/Single High Risk Area)	36	0.900
H	(Only Use/Multiple High Risk Area)	39	0.975

This new table is used to compute p(Effects) for Example 3 and Example 8, respectively. The conditions for each example are contained in Table F4-4. Figures F4-3 and F4-4 illustrate the results; the purple diamonds on each figure are the M&S Use Risks for p(Effects) levels A through H, bottom to top of graph.

Table F4-4: Examples of p(Effects) Calculation

Example	CLARITY C_1 Level [$p(C_1)$]	IMPORTANCE C_2 Level [$p(C_2)$]	CONFIDENCE C_3 Level [$p(C_3)$]	p(Causes) [$p(C_1 \cup C_2 \cup C_3)$]	p(Effects)	M&S Use Risk
3	Unclear; A [0.833]	Grave/Partial; F [0.846]	Very high; A [0.05]	0.976	0.025	0.001; Very low
					0.100	0.012; Very low
					0.225	0.053; Low
					0.400	0.162; Medium
					0.600	0.357; Very high
					0.775	0.590; Very high
					0.900	0.793; Very high
					0.975	0.928; Very high
8	Lucid; A [0.167]	Serious/Complete; D [0.423]	Very high; A [0.05]	0.543	0.025	0.006; Very low
					0.100	0.028; Very low
					0.225	0.071; Low
					0.400	0.146; Medium
					0.600	0.255; High
					0.775	0.369; Very high
					0.900	0.462; Very high
					0.975	0.522; Very high

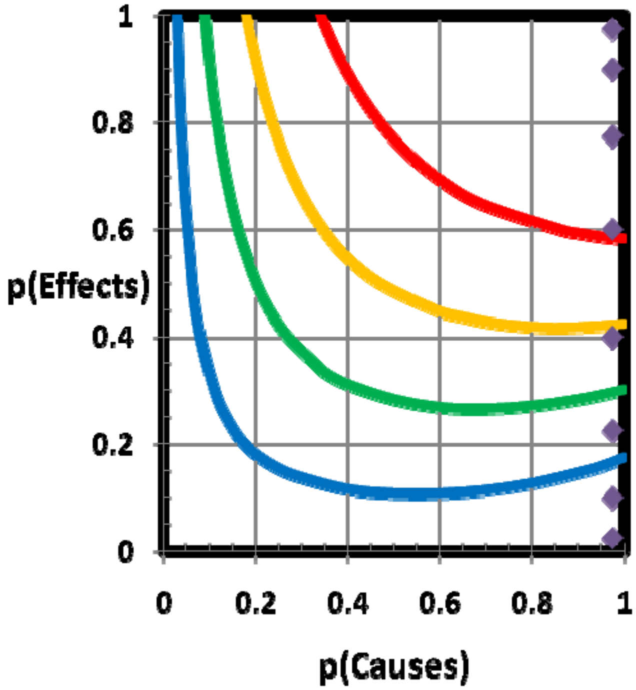


Figure F4-3: Example 3 Using New p(Effects)

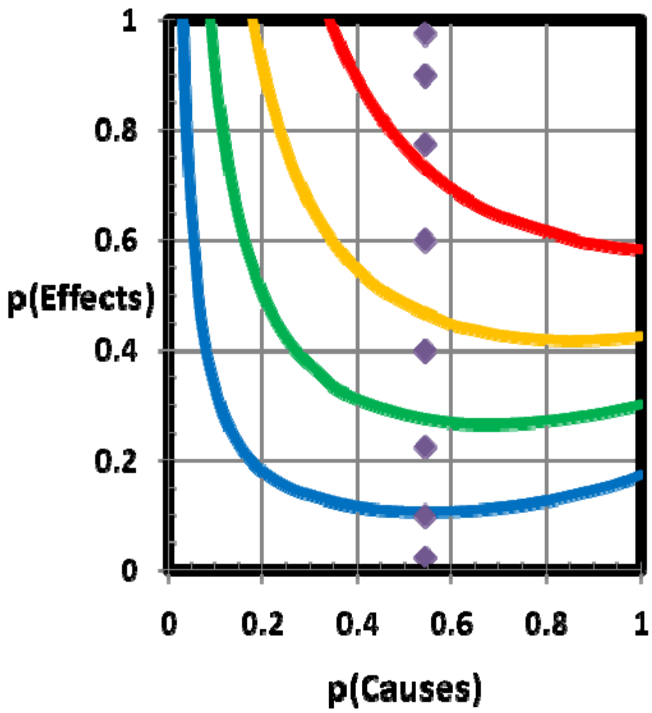


Figure F4-4: Example 8 Using New p(Effects)

APPENDIX F5: ALTERNATE p(C_1) ANALYSIS

The analysis presented in the previous sections employs a placeholder state table for the C_1 factor, Clarity of Intended Use. Table F5-1 duplicates that placeholder state table comprised of three levels with no further details:

Table F5-1: State Table for Factor C_1 (Clarity), Assignment of p(C_1)

Factor Level	Clarity of Intended Use	Level Weighting	p(C_1)
A	Lucid	1	0.167
B	Partial clarity	3	0.5
C	Unclear	5	0.833

This table sufficed to illustrate how the contribution of the Clarity of Intended Use affects the M&S Use Risk. In this appendix, an alternate formulation for determining p(C_1) is illustrated. This alternate approach considers the factor as the sum of the effects of four sub-factors, thus giving a finer granularity to the definition of C_1. A sub-factor state table is a function of the logical atoms encompassing its definition. The value for p(C_1) is derived from the cumulative distribution of a weighted raw score formed by the evaluation of the four sub-factors (similar to the approach for establishing the C_3 factor; see, for example, Figure F-2).

The sub-factors are as follows: C_{11} is "INTENDED USE"; C_{12} is "M&S REQUIREMENTS; C_{13} is "ACCEPTABILITY CRITERIA"; and, C_{14} is "REFERENT". Identification of the specific atoms and descriptions of the sub-factor state tables follow.

STATE TABLE FOR SUB-FACTOR C_{11}: INTENDED USE

For the INTENDED USE sub-factor, Table F5-2 provides a 3-level state table:

Table F5-2: State Table for INTENDED USE, Sub-factor C_{11}

Level	Sentence	Weighting
A	$x \wedge y$	5
B	$x \wedge \sim y$	3
C	$\sim x \wedge \sim y$	1

where, $x \equiv$ statement of intended use(s) exists, and $y \equiv$ a relationship of the capability to the intended use(s) is established. Note that of the four combinations formed by the adjunct of these two atoms, one combination, $\sim x \wedge y$, is outside the state space; it is determined that a relationship between capability and an intended use without having a statement of intended use cannot occur.

STATE TABLE FOR SUB-FACTOR C_{12}: M&S REQUIREMENTS

For the M&S REQUIREMENTS sub-factor, Table F5-3 provides a 3-level state table:

Table F5-3: State Table for M&S REQUIREMENTS, Sub-factor C_{12}

Level	Sentence	Weighting
A	$x \wedge y$	5
B	$x \wedge \sim y$	3
C	$\sim x \wedge \sim y$	1

where $x \equiv$ statement of the requirement exists, and $y \equiv$ the statistics for the requirement are established. Note that of the four combinations formed by the adjunct of these two atoms, one combination, $\sim x \wedge y$, is outside the state space; it is determined that the statistics for a requirement cannot be established unless there is a statement of the requirement.

STATE TABLE FOR SUB-FACTOR C_{13}: ACCEPTABILITY CRITERIA

For the ACCEPTABILITY CRITERIA sub-factor, Table F5-4 provides a 3-level state table:

Table F5-4: State Table for ACCEPTABILITY CRITERIA, Sub-Factor C_{13}

Level	Sentence	Weighting
A	$x \wedge y$	5
B	$x \wedge {\sim}y$	3
C	${\sim}x \wedge {\sim}y$	1

where x ≡ statement of the ACCEPTABILITY CRITERIA exists, and y ≡ the statistics for the acceptability criteria are established. Note that of the four combinations formed by the adjunct of these two atoms, one combination, ${\sim}x \wedge y$, is outside the state space; it is determined that the statistics for the acceptability criteria cannot be established unless there is a statement of the requirement.

STATE TABLE FOR SUB-FACTOR C_{14}: REFERENT

For the REFERENT sub-factor, Table F5-5 provides a 4-level state table:

Table F5-5: State Table for REFERENT, Sub-Factor C_{14}

Level	Sentence	Weighting
A	$x \wedge y \wedge z$	9
B	$(x \wedge {\sim}y \wedge z) \vee (x \wedge y \wedge {\sim}z)$	6
C	$x \wedge {\sim}y \wedge {\sim}z$	3
D	${\sim}x \wedge {\sim}y \wedge {\sim}z$	1

where x ≡ referent identified, y ≡ referent source is established, and z ≡ referent's statistical descriptions are established. Note that of the eight combinations formed by the adjunct of these three atoms, three combinations are not in the state space. If the referent is unidentified then it is not possible to establish a referent's source or statistics.

Table F5-6 summarizes the weightings of the levels of the four sub-factors before adjustment:

Table F5-6: Summary of Sub-Factor Weightings before Adjustment

Level	Sub-Factor C_{11}	Sub-Factor C_{12}	Sub-Factor C_{13}	Sub-Factor C_{14}
A	5	5	5	9
B	3	3	3	6
C	1	1	1	3
D	-	-	-	1
SUM:	*9*	*9*	*9*	*19*

Because the numbers of levels in all sub-factors are not equal, a multiplication adjustment insures each sub-factor has, in total, equivalent influence on C_1. The weightings for sub-factors C_{11}, C_{12}, and C_{13} are multiplied by 19 (the sum of the level weightings for sub-factor C_{14}), and the weighting for sub-factor C_{14} is multiplied by 9 (the sum of the level weightings in each of the other three sub-factors). Table F5-7 represents the adjusted state tables:

Table F5-7: Adjusted Sub-Factor Weightings

Level	Sub-Factor C_{11}	Sub-Factor C_{12}	Sub-Factor C_{13}	Sub-Factor C_{14}
A	95	95	95	81
B	57	57	57	54
C	19	19	19	27
D	-	-	-	9
SUM:	*171*	*171*	*171*	*171*

A weighted raw score established from the four sub-factors is as follows:

$$\text{WRS} = [\text{Sub-factor } C_{11} \text{ ADJ Weighting}] + [\text{Sub-factor } C_{12} \text{ ADJ Weighting}]$$

$$+[\text{Sub-factor } C_{13} \text{ ADJ Weighting}] + [\text{Sub-factor } C_{14} \text{ ADJ Weighting}]$$

There are 108 combinations for the WRS, but there are only 28 unique values for the score. Figure F5-1 illustrates the cumulative distribution in percentile form.

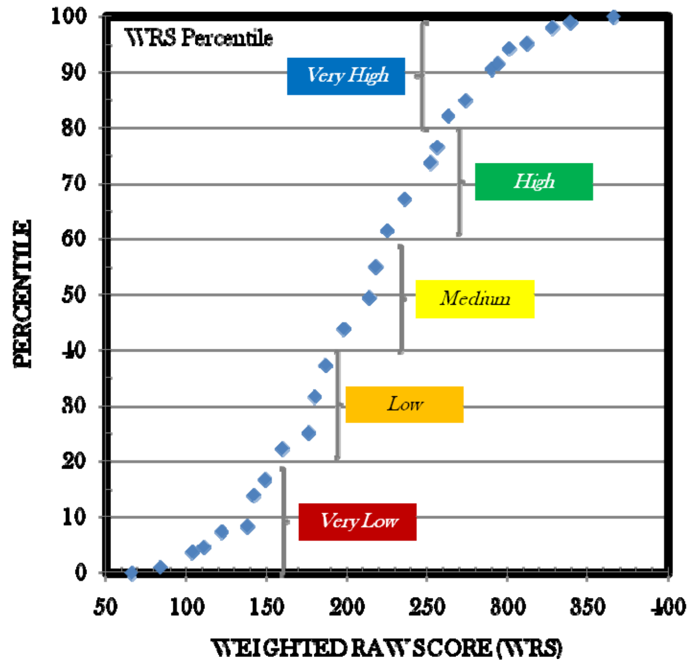


Figure F5-1: Weighted Raw Score Percentiles for Factor C_1 (Clarity)

Establishing the Factor C_1 State Table

If, as suggested, that the result for CLARITY OF INTENDED USE be reported in quintiles (i.e., percentiles 0-20, 20-40, etc.), then the intervals will be grouped into five separate levels. A state table leading to an assignment of $p(C_1)$, Table F5-8, may be built around the results of Figure F5-1 (and the maximum information entropy arguments of Appendix F2).

Table F5-8: State Table for Factor C_1 (Clarity), Assignment of $p(C_1)$

Factor Level	Clarity of Intended Use	Level Weighting	$p(C_1)$
A	WRS percentiles 80 to ≤100; Very high	1	0.1
B	WRS percentiles 60 to <80; High	3	0.3
C	WRS percentiles 40 to <60; Medium	5	0.5
D	WRS percentiles 20 to <40; Low	7	0.7
E	WRS percentiles 0 to <20; Very low	9	0.9

The maximum information entropy principle determines the weights of each quintile by knowing only the order of the levels. The values for $p(C_1)$ from this table could be used in lieu of those of Table F-2 if more resolution of the clarity of the indented use is desired.

APPENDIX G: GLOSSARY

G.1 ABBREVIATIONS AND ACRONYMS

C

C_1	Cause 1 - Lack of clarity of intended use leading to misuse (i.e., Clarity)
C_2	Cause 2 - Adverse impact on decision if capability is not achieved (i.e., Importance)
C_3	Cause 3 - Incorrect recommendation to employ or not to employ M&S Results relative to that capability (i.e., Confidence)
CM	Conceptual Model

D

DoD	Department of Defense
DP	Development Product

J

JHU/APL	The Johns Hopkins University Applied Physics Laboratory

M

M&S	model(s), simulation(s), and associated data
M&S CO	Modeling and Simulation Coordination Office
M&S SC	Modeling and Simulation Steering Committee
MIE	Maximum Information Entropy
MIL-STD	Military Standard
MOE	Measure of Effectiveness
MOP	Measure of Performance
MURM	M&S Use Risk Methodology

R

RBA	Risk-Based Verification, Validation, and Accreditation
RF	Radio Frequency
RPG	Recommended Practices Guide

S

SME	Subject Matter Expert

U

UR	Use Risk
URS	Unweighted Raw Score

V

V&V	Verification and Validation
VCM	Verification and Validation Composite Model
VPMM	Verification and Validation Process Maturity Model
VV&A	Verification, Validation, and Accreditation

W

WRS	Weighted raw score

G.2 DEFINITIONS

For the purposes of this document, the following terms and definitions apply.

Acceptability Criteria. A set of standards that a particular model or simulation and its associated data must meet to be accredited for a specific purpose.

Accreditation. The official certification that a model or simulation and its associated data are acceptable for use for a specific purpose. [Reference (s)]

Accreditation Agent. The organization designated to conduct an accreditation assessment for an M&S application. [Reference (g)]

Accreditation Authority. The organization or individual responsible to approve the use of models, simulations, and their associated data for a particular application. [Reference (s)]

Data Verification and Validation. The process of verifying the internal consistency and correctness of data and validating that it represents real-world entities appropriate for its intended purpose or an expected range of purposes. [Reference (s)]

Intended Use Statement. A description of the problem to be addressed by a model or simulation and its associated data, including the system or process being represented and the role it plays in the overall program. [Reference (g)]

M&S Developer. The agency that develops an M&S or the agency that is overseeing the M&S development by a contractor. [Reference (g)]

M&S Program Manager. The individual responsible for planning and managing resources for simulation development, directing the overall simulation effort, and overseeing configuration management and maintenance of the simulation. In legacy simulation reuse when a major modification effort is involved, the M&S User may designate an M&S Program Manager to plan and manage the modification effort. [Reference (g)]

M&S Proponent. The DoD Component organization that has primary responsibility to initiate development and life-cycle management of the reference version of one or more models and/or simulations. [Reference (g)]

M&S Requirements. Conditions or capabilities that must be met or possessed by a model, simulation and its associated data to satisfy a contract, standard, specification, program document, other formal document, or for regulatory compliance. There may be technical, non-technical, system-level, product-level, component-level, and/or stakeholder requirements for the development, modification, or use of the M&S.

M&S Use Risk. The probability that inappropriate application of M&S Results for the intended use will produce unacceptable consequences to the decision-maker.

M&S User. M&S User is the term used to represent the organization, group, or person responsible for the overall application. The M&S User needs to solve a problem or make a decision and wants to use modeling or simulation to do so. The M&S User defines the requirements, establishes the criteria by which model or simulation fitness will be assessed, determines what method or methods to use, makes the accreditation decision, and ultimately accepts the results. [Reference (g)]

Measure of Effectiveness (MOE). A measure designed to correspond to the M&S Intended Use, M&S Requirements, Acceptability Criteria, and achievement of desired results when applying M&S..

Measure of Performance (MOP). A measure of M&S capability performance in a given environment expressed as distinctly quantifiable performance features. MOPs define the specific performance issues identified by the Measures of Effectiveness (MOEs).

Model. A physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process. [Reference (s)]

Referent. A codified body of knowledge about a thing being simulated.

Simuland. The system being simulated by a simulation. [Reference (f)]

Simulation. A method for implementing a model over time. [Reference (s)]

Stakeholder. Anyone who might influence the priority and importance of a particular M&S capability.

Validation. The process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model. [Reference (s)]

Verification. The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications. [Reference (s)]

VV&A practitioner. The organization, group, or person implementing VV&A processes.

G.3 MATHEMATICAL LOGIC SYMBOLS AND TERMINOLOGY

For the purposes of this document, the following symbols and definitions apply.

$\cup$	union
$\cap$	intersection
$\wedge$	logical conjunction (i.e., and)
$\vee$	logical disjunction, "or"
$\Rightarrow$	implication, "if, then"

$\equiv$	equivalent
$=$	equality, is equal to; equals
$+$	addition, plus, add
—	subtraction, minus, subtract
x	multiplication, times
~	inverse relation, "not"
Σ	summation
p(x)	probability (i.e., value of "x being true")
Causes	the union of all inappropriate applications of M&S Results for the intended use is true
Effects	the unacceptable consequences to the decision-maker are true
p(Causes)	the probability of the causes of inappropriate application
p(Effects)	the probability of the effects of unacceptable consequences to the decision-maker

$S = \Sigma\, p_i \log_2 [1/p_i]$ equation for information entropy

S is information entropy

p_i is probability for the i^{th} class.

S_{max} the maximum information entropy of a system

This page intentionally left blank.

NATIONAL SECURITY ANALYSIS DEPARTMENT
THE JOHNS HOPKINS UNIVERSITY • APPLIED PHYSICS LABORATORY
Johns Hopkins Road, Laurel, Maryland 20723-6099