NDIA Conference: Physics-based Modeling in Design and Development for U.S. Defense, November 14-17, 2011 (Denver, Colorado)

A Tutorial on the Identification of M&S Uncertainty and Assessment of M&S Use Risk

Simone Youngblood (JHU/APL) -- Presenter
Peter Pandolfini (JHU/APL)
Dale Pace (Consultant)

APL
The Johns Hopkins University
APPLIED PHYSICS LABORATORY

Tutorial Addresses Two Issues

1) Assessment of M&S Use Risk using the M&S Use Risk Methodology (MURM)
   - Why MURM was developed
   - What MURM is, an overview
   - Basic steps in MURM
   - MURM related calculations
   - Expected MURM outputs and benefits
2) Incomplete consideration of uncertainty in M&S
   - The Problem: Wrong conclusions can be drawn from M&S results
   - Suggested Solution: Use a paradigm that facilitates comprehensive consideration of M&S uncertainty

MURM Background

MURM Project Objectives

Leverage existing concepts to evolve a methodology to:

- **Weigh VV&A investments against the risk of making a bad decision due to unreliable M&S results**
- **Tailor the V&V and Accreditation Efforts based on risk**
- **Perform Methods/Technique/Resource Trade-offs**

Why MURM?

Previous risk-based M&S assessments have deficiencies:

- Lack cogent mathematical foundation
- Sometimes included unintended bias
- Can't explicitly relate V&V endeavors to risk

Fundamental Driver for MURM: the need to combine M&S risk assessment and VV&A planning in a coherent fashion.

M&S Use Risk Methodology (MURM) provides:

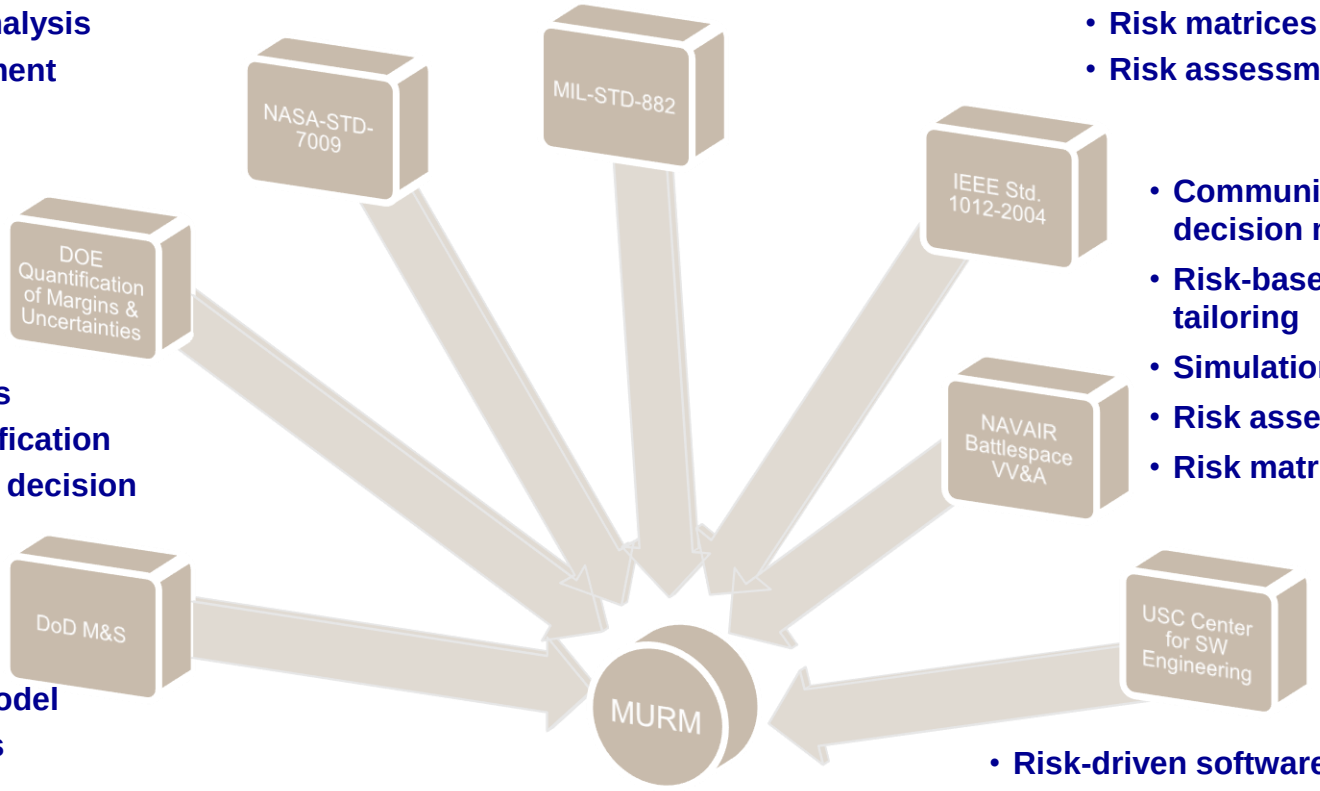
- Coherent math foundation for M&S Use Risk
- Minimizes or avoids unintended bias
- Explicit relation of V&V to M&S Use Risk
- Facilitates automation of M&S Use Risk assessment

The MURM Builds upon Existing Concepts

- Communication to decision makers
- Uncertainty quantification
- Credibility assessment
- Sensitivity analysis
- Risk assessment

- Severity categories
- Probability levels
- Risk assessment values
- Risk acceptance levels

- Risk-based V&V tailoring
- Software integrity level
- Risk matrices
- Risk assessment



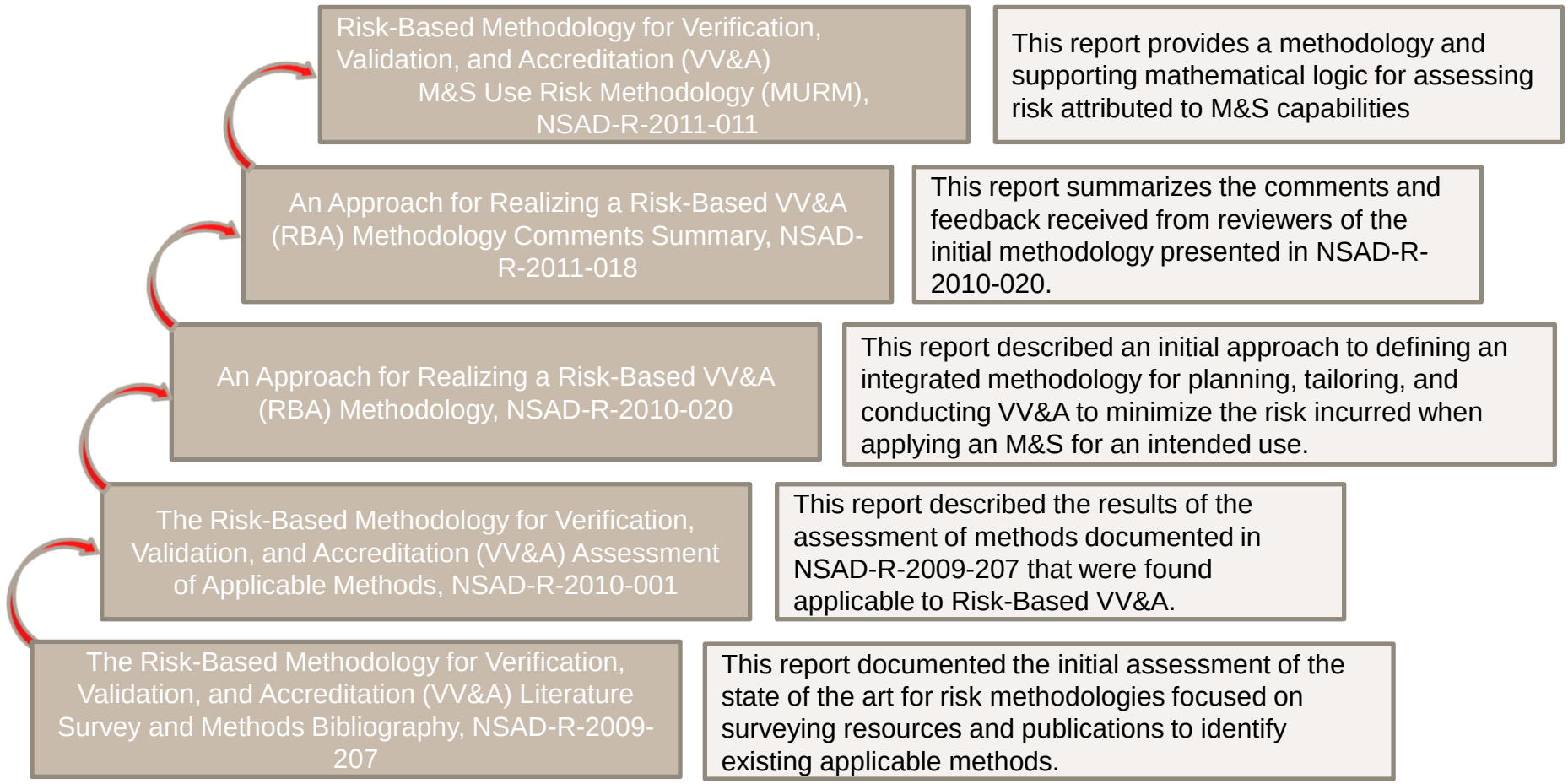
- Confidence ratios
- Sensitivity analysis
- Uncertainty quantification
- Communication to decision makers

- Communication to decision makers
- Risk-based VV&A tailoring
- Simulation importance
- Risk assessment
- Risk matrices

- V&V Composite Model
- Validation Process Maturity

- Risk-driven software development
- Spiral development model

Risk-Based Methodology for VV&A Product Overview



Risk & Risk-related Analysis Is Complex

- DAU review of standards found half-dozen definitions each with negative, neutral, and broad connotations for risk, which indicates the complexity of the subject.
- Many simplify risk and treat the approach as if it fully addresses risk. Common definition is:

  Risk = (likelihood of Error) * (Consequence of Error)

- This definition mixes a probability with a non-probability, resulting in an expected value whose dimensions change from one assessment to another.
- MURM takes a different approach.

Many Aspects of M&S-related Risk

- Descriptors often indicate the aspect of risk being addressed: e.g., programmatic risk, technical risk, operational risk, etc.
- MURM focuses on a definition that accommodates both assessment of the consequences of using M&S results and the impact of V&V planning and execution.
- MURM is expressed first in words and then in set theory so that mathematical coherence may be obtained for MURM.

MURM Definition

M&S Use Risk – Key Definition

- MURM is not just a collection of tables and figures. MURM applies an underlying mathematical formula based on the definition 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.

- For unambiguous understanding of this definition of M&S Use Risk, two important conditions are noted:
  (1) both the inappropriate application and the unacceptable consequences occur
  (2) the unacceptable consequences are implied by the inappropriate application.
- M&S Use Risk is a probability & whole probability space is addressed
- Math Logic foundation enables:
  - explicit relationship of M&S Use Risk to V&V endeavors
  - facilitates automation of UR assessment

M&S Use Risk Mathematical Form

To put the definition into a mathematical form that can be used in a numerical calculation of M&S Use Risk, the definition is 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:

> **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:

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

M&S Use Risk Mathematical Form (cont.)

Where

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

With

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

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

$C_3 \equiv$ 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 of unacceptable consequences to the decision-maker. p(Effects) is based on M&S Impact and M&S Reliance.

M&S USE RISK DEFINITION Evolution

IN WORDS:

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

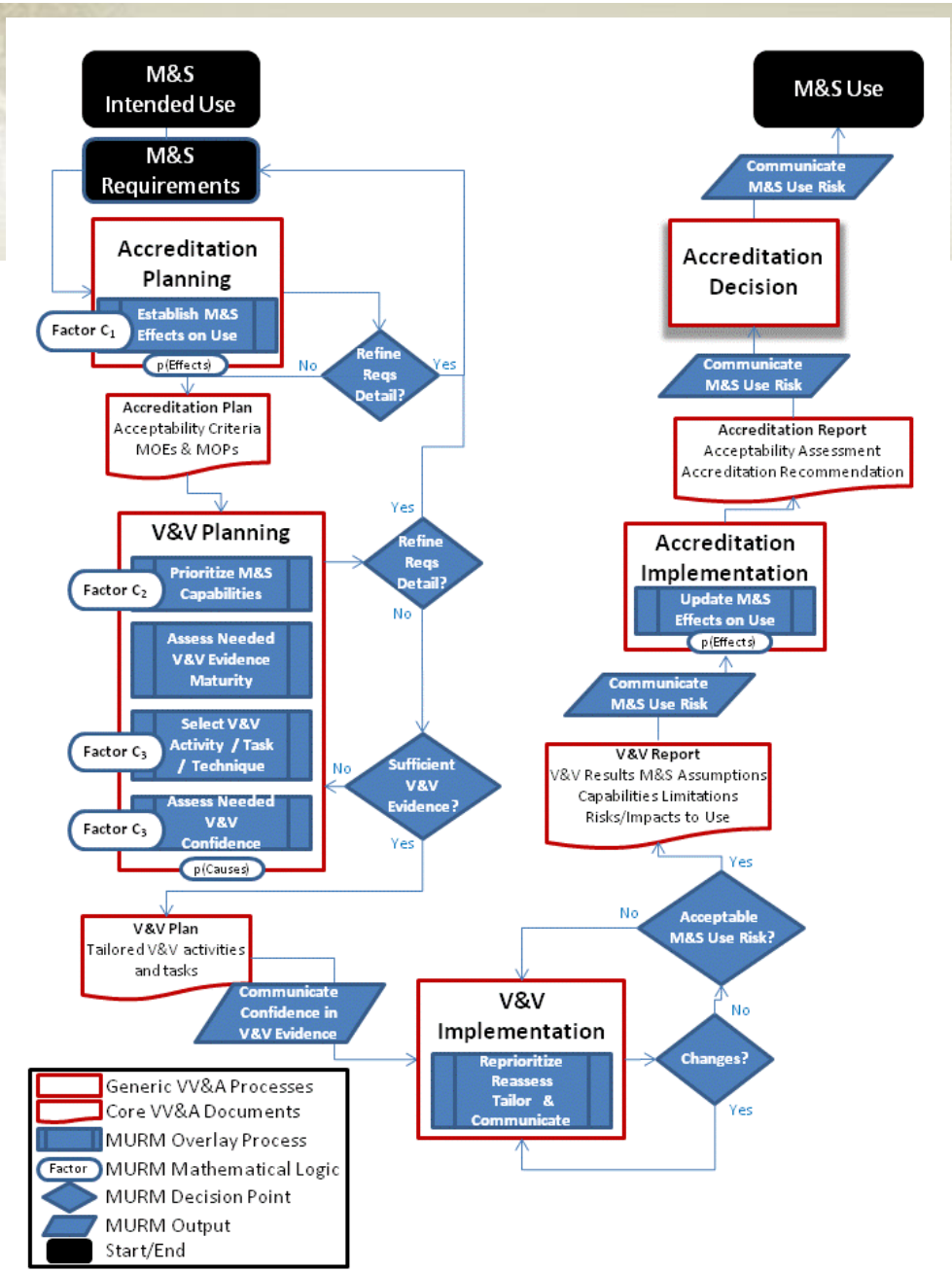
IN MATHEMATICAL LOGIC:

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

ALGEBRAIC FORM for PROBABILITY CALCULATIONS:

M&S Use Risk = p(Causes) x p(Effects) x [1- p(Causes) + p(Causes) x p(Effects)]

APL



MURM Processes & Relation to VV&A Activities

For this tutorial, four basic steps are shown in the next slide to describe and explain how to execute MURM

APL

Steps in the RBA Methodology

Determine Needed Capabilities

- Obtain requirements from users & their representatives
- Identify needed representational capabilities
- Organize needed capabilities into related groupings

Characterize Use Risk Areas

- Determine if the user can specify the M&S role
- Determine M&S importance to each capability
- Determine needed validation maturity level from assigned importance

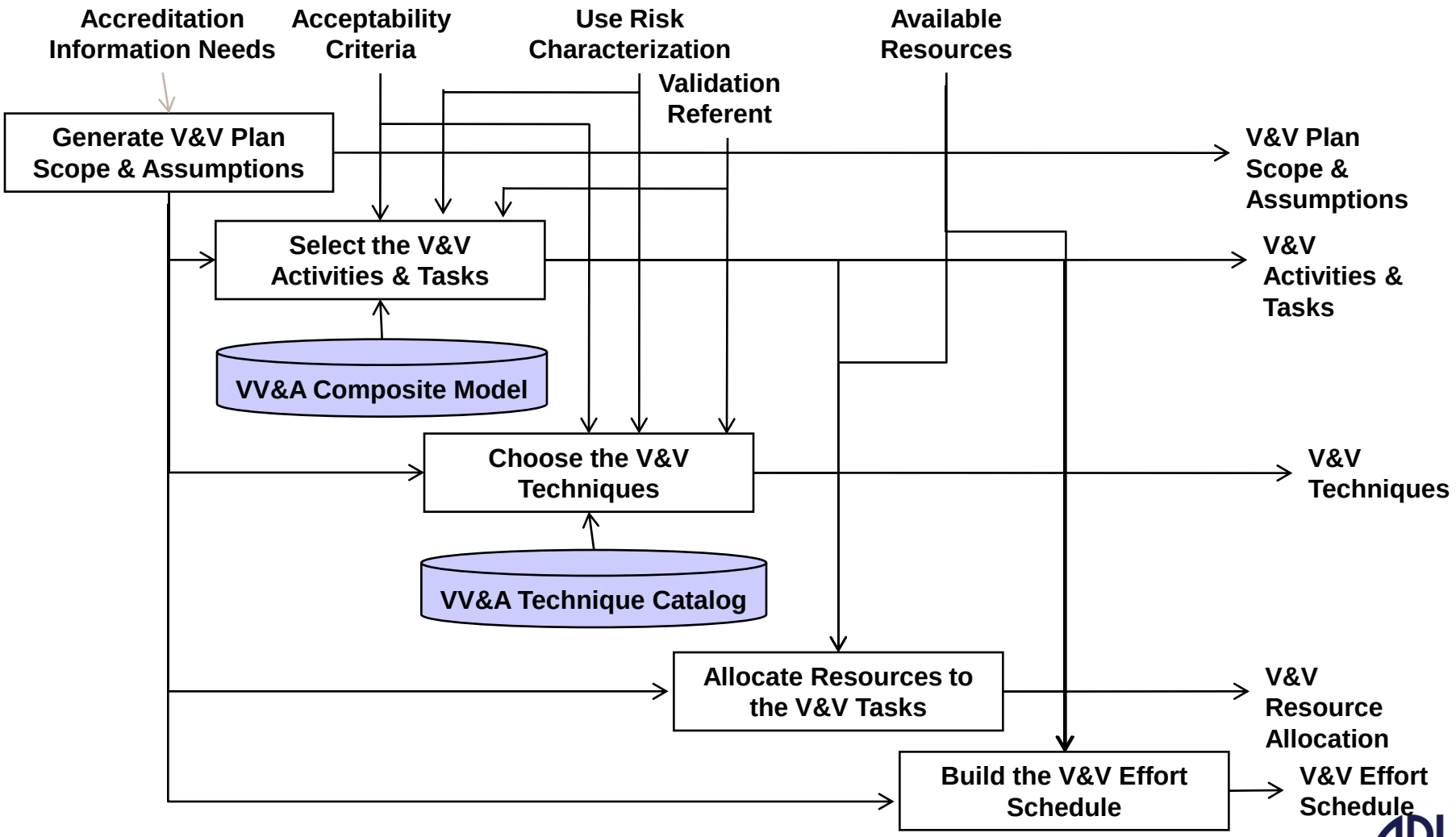
Plan V&V Effort

- Choose V&V activities & tasks to be performed
- Choose V&V techniques to meet needed validation maturity
- Estimate schedule, costs & resource requirements

Collect V&V Evidence

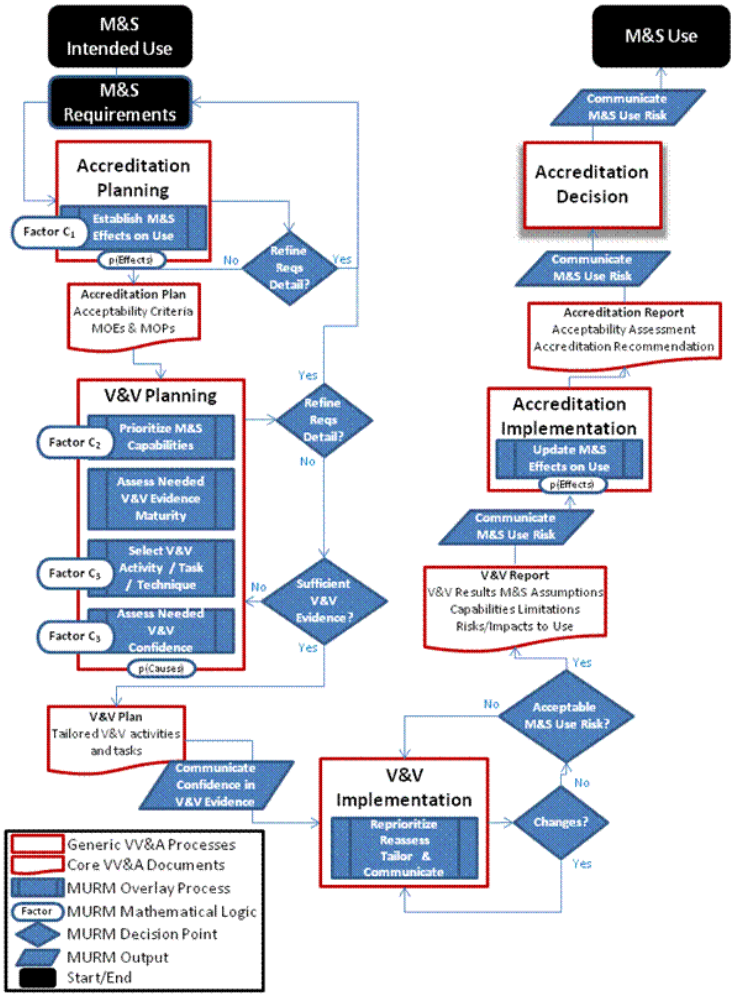
- Characterize M&S capabilities
- Characterize M&S limitations
- Communicate V&V evidence to the M&S users

The RBA Tailoring Concept



MURM Basic Steps

PART IV: MURM'S BASIC STEPS



(1) Establish intended use(s)
(2) Enumerate requirements
(3) Prioritize requirements
(4) Establish/select "Causes" state tables (C_1, C_2, C_3)
(5) Establish initial V&V plan
(6) Evaluate, requirement-by-requirement state levels for C_1, C_2, C_3
(7) Compute p(Causes)
(8) Establish/select "Effect" state table
(9) Evaluate, requirement-by-requirement state levels for Effects; compute p(Effects)
(10) Compute M&S User Risk for each requirement (f(p(Causes),p(Effects))
(11) Evaluate acceptability of each requirement's M&S User Risk
(12) Accept or modify V&V plan (return to step (5))

APL

Decomposition of P(Causes)

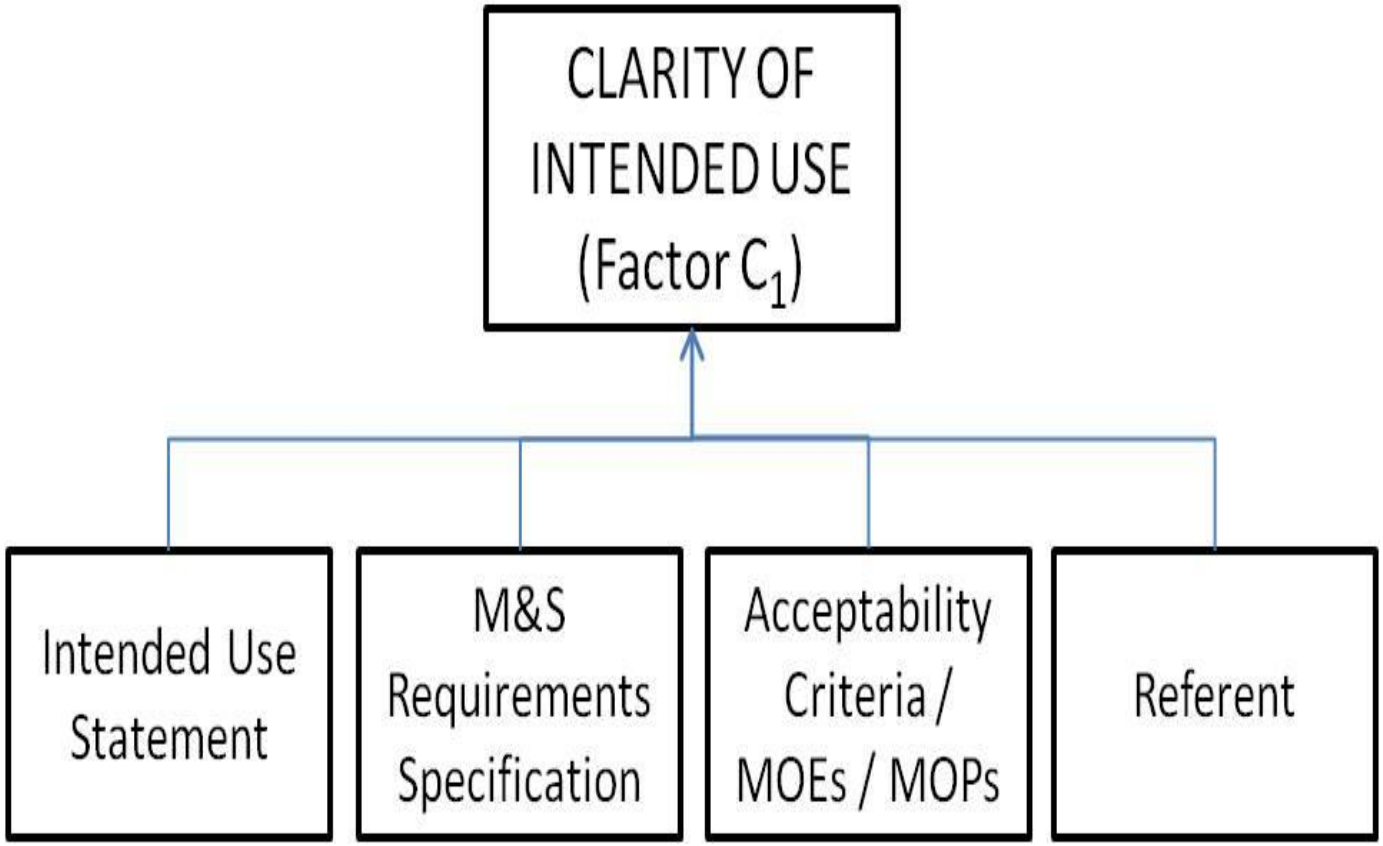
$$p(\text{Causes}) = p(C_1 \cup C_2 \cup C_3)$$

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

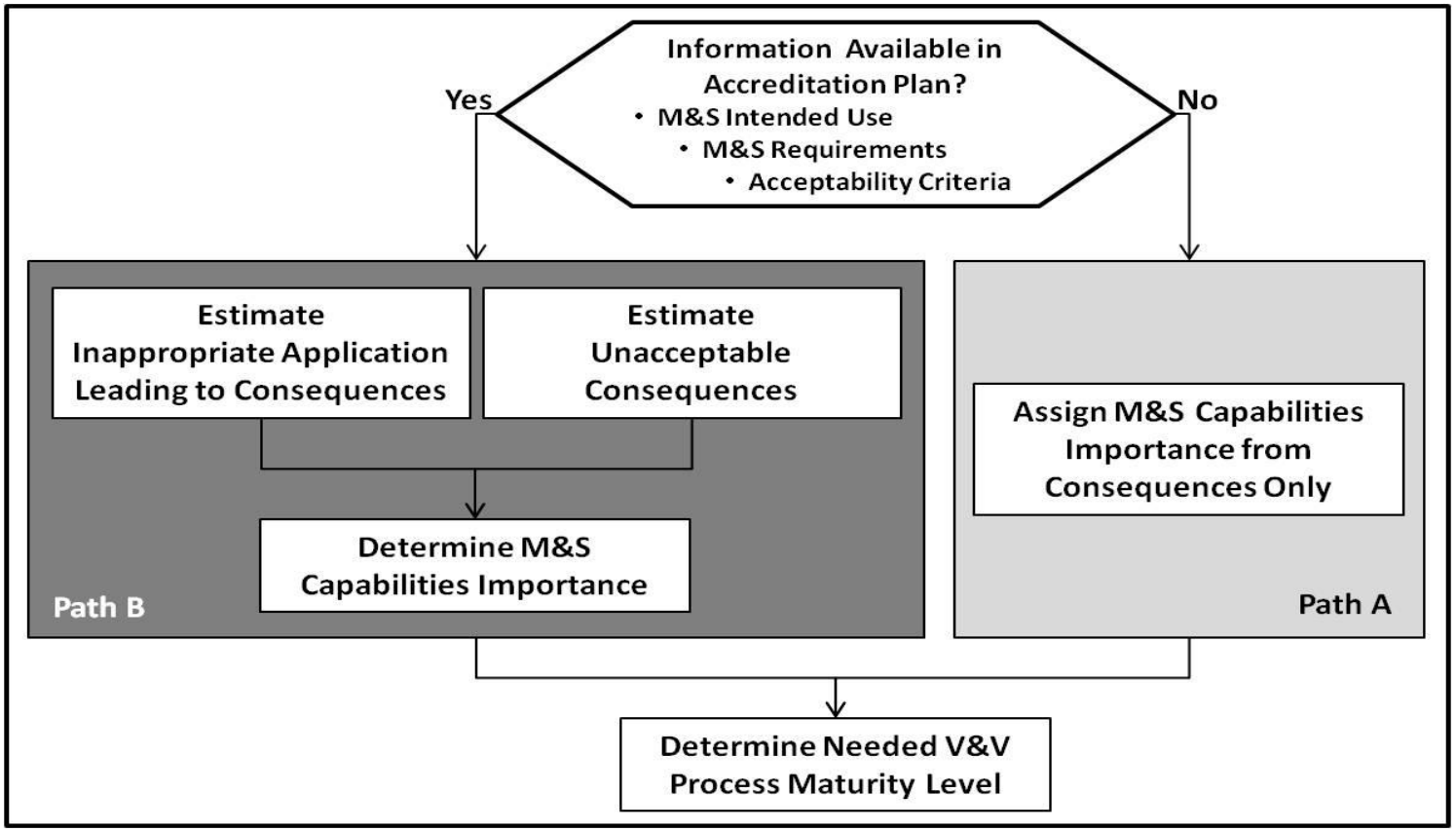
$C_2 \equiv$ Adverse impact on decision if capability is not achieved (Importance), and

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

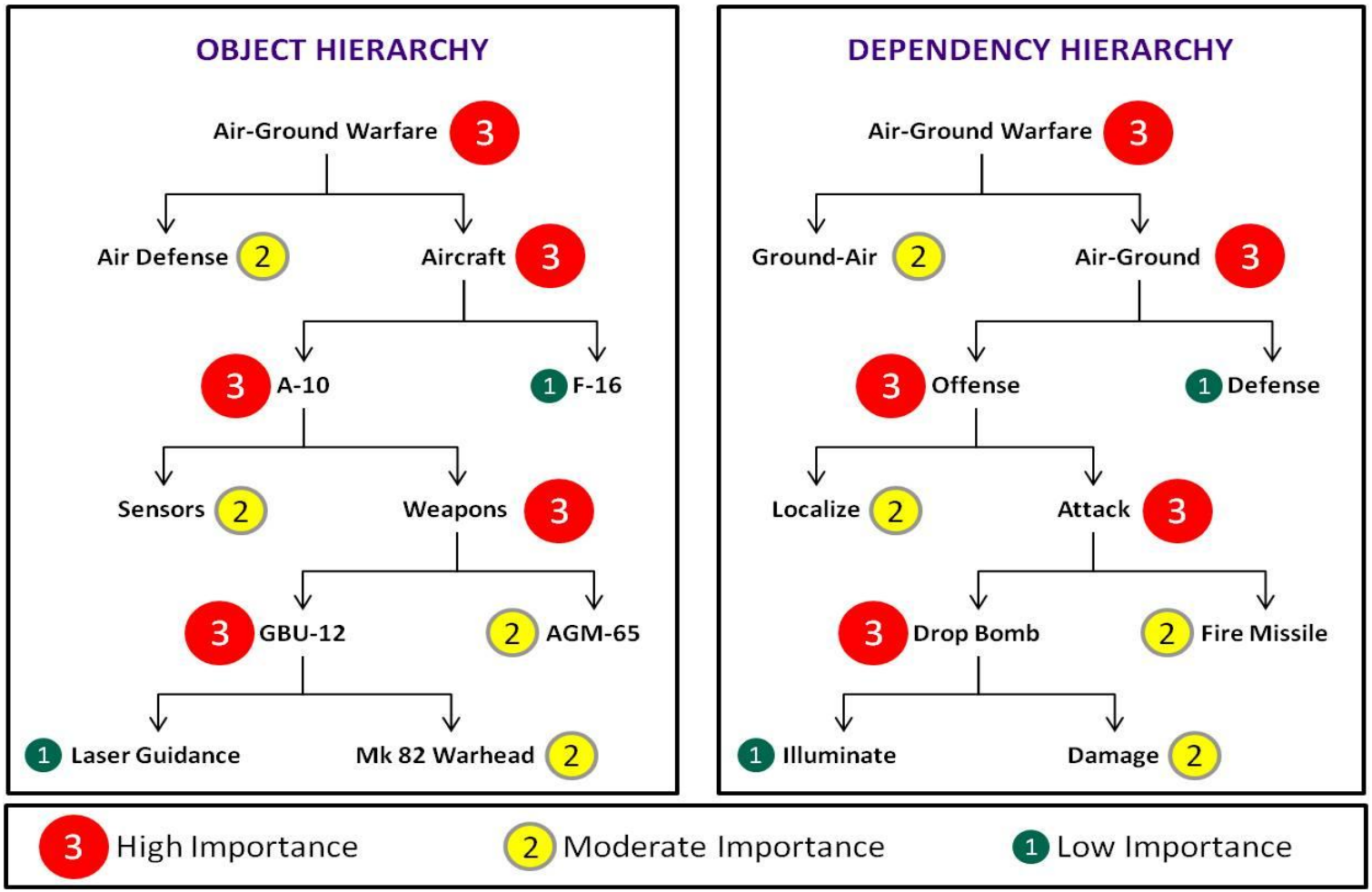
CLARITY OF INTENDED USE (MURM FACTOR C_1)



Assigning Importance to M&S Capabilities (MURM Factor C_2)



ASSIGNING IMPORTANCE TO M&S CAPABILITIES (MURM FACTOR C_2)



Example: State Table Probabilities For Causes

Table F-2: 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

Examples of factor state tables

Built using mathematical logic & maximum information entropy principle

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

APL

CONFIDENCE (MURM FACTOR C_3)

<table>
<tr><td colspan="6">V&V Technique: Desk Checking / Self-Inspection</td></tr>
<tr><td colspan="2">Class:
▪ Informal</td><td colspan="2">V&V Category:
▪ Verification</td><td colspan="2">M&S Phases:
▪ M&S Requirements
▪ Conceptual Model
▪ M&S Design
▪ M&S Development</td></tr>
<tr><td colspan="6">Abstract: Refers to a technical team/peer examination of an M&S artifact.</td></tr>
<tr><td colspan="6">Objective: Used to ensure correctness, completeness, consistency, and clarity.</td></tr>
<tr><td colspan="6">Examination: Conducts syntax review, cross-reference examination, convention violation assessment, detailed comparison to specifications, code reading, control flow-graph analysis, and path sensitizing.</td></tr>
<tr><td colspan="3">Inputs:
▪ Source code
▪ Input/output data</td><td colspan="3">Outputs:
▪ Error discovery
▪ Recommended corrections</td></tr>
<tr><td colspan="3">Level of Effort:</td><td colspan="3">Rigor:</td></tr>
<tr><td>High</td><td>Medium</td><td>Low</td><td>High</td><td>Medium</td><td>Low</td></tr>
<tr><td colspan="6">References:
▪ Beizer, B., Software Testing Techniques (2nd ed.), Van Nostrand Reinhold, New York, 1990.
▪ Hayardeny, A., Fienblit, S., & Farchi, E, Distributed desk checking, Concurrency and Computation: Practice and Experience, 19, pp. 295-309, 2007.</td></tr>
</table>

Decomposition of P(Effects)

p(Effects) is the probability of the effects resulting from unacceptable consequences to the decision-maker and is derived from to key components: M&S Impact and M&S Reliance

- **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.**

Example: State Table Probabilities For Effects

Example of simple Effects state table

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

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.

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.

Example: State Table Probabilities For Effects

Example of more complex Effects state table

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

APL

EXAMPLES
CALCULATIONS OF P(CAUSES)

Require-ment No.	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

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

APL

EXAMPLES
CALCULATIONS OF M&S USE RISK

M&S Use Risk = p(Causes) p(Effects) [1- p(Causes) + p(Causes) p(Effects)]

Suppose for Requirement #2, it is determined Table F4-3 Effects Level is "A"

p(Effects) = 0.025 (from Table F4-3)
p(Causes) = 0.930 (from previous slide)

M&S Use Risk = 0.930×0.025 × [1- 0.930 + 0.930×0.025] = 0.002 (*Very Low*)

Suppose for Requirement #4, it is determined Table F4-3 Effects Level is "C"

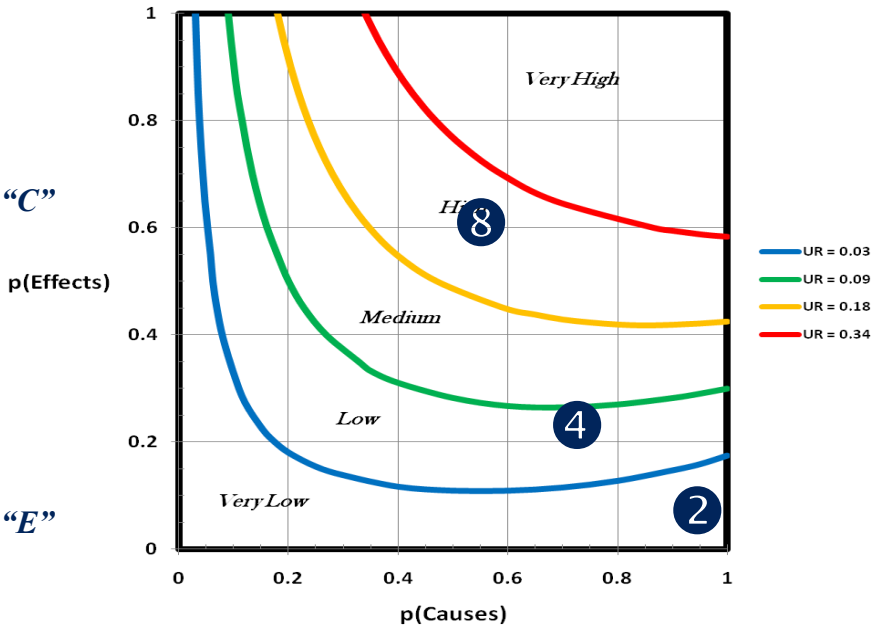
p(Effects) = 0.225 (from Table F4-3)
p(Causes) = 0.735 (from previous slide)

M&S Use Risk = 0.735×0.225 × [1- 0.735 + 0.735×0.225] = 0.071 (*Very Low*)

Suppose for Requirement #8, it is determined Table F4-3 Effects Level is "E"

p(Effects) = 0.600 (from Table F4-3)
p(Causes) = 0.543 (from previous slide)

M&S Use Risk = 0.543×0.600 × [1- 0.543 + 0.543×0.600] = 0.255 (*High*)



EXAMPLES
CALCULATIONS OF M&S USE RISK

M&S Use Risk = p(Causes) p(Effects) [1- p(Causes) + p(Causes) p(Effects)]

Suppose for Requirement #2, it is determined Table F4-3 Effects Level is "A"

p(Effects) = 0.025 (from Table F4-3)
p(Causes) = 0.930 (from previous slide)

M&S Use Risk = 0.930×0.025 × [1- 0.930 + 0.930×0.025] = 0.002 (*Very Low*)

Suppose for Requirement #4, it is determined Table F4-3 Effects Level is "C"

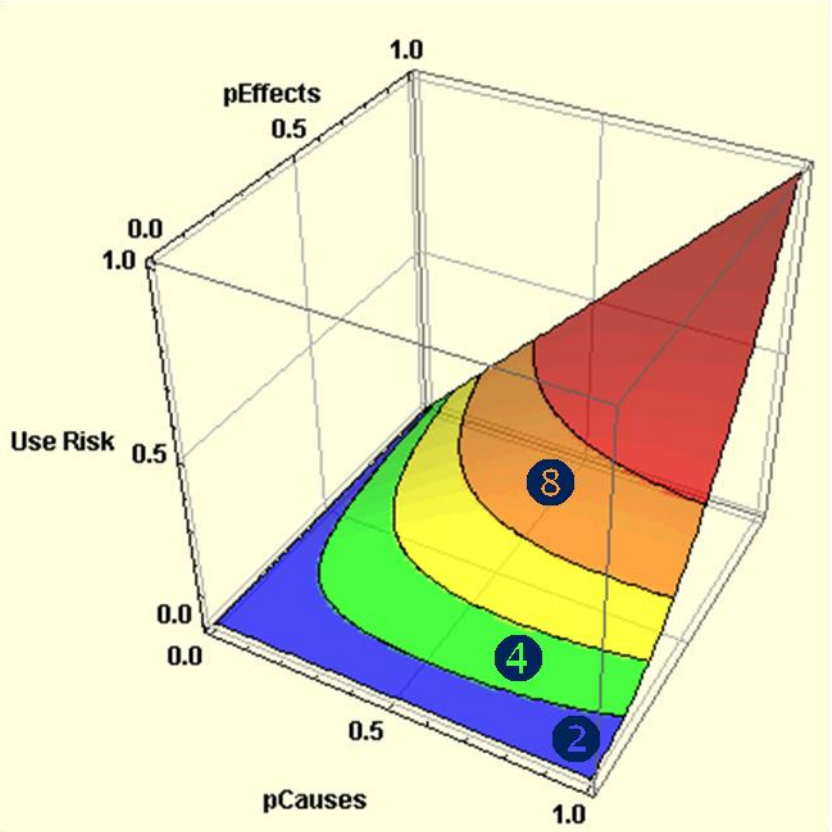
p(Effects) = 0.225 (from Table F4-3)
p(Causes) = 0.735 (from previous slide)

M&S Use Risk = 0.735×0.225 × [1- 0.735 + 0.735×0.225] = 0.071 (*Low*)

Suppose for Requirement #8, it is determined Table F4-3 Effects Level is "E"

p(Effects) = 0.600 (from Table F4-3)
p(Causes) = 0.543 (from previous slide)

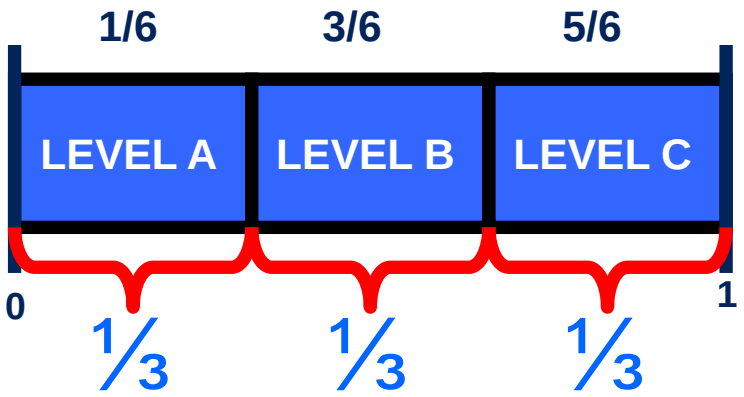
M&S Use Risk = 0.543×0.600 × [1- 0.543 + 0.543×0.600] = 0.255 (*High*)



EXAMPLE PROBABILITIES FROM STATE TABLES

Table F-2: 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



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

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

Built using mathematical logic & maximum information entropy principle

CENTROIDS OF INTERVALS AT 0.167, 0.500, and 0.833 or a ratio of 1:3:5

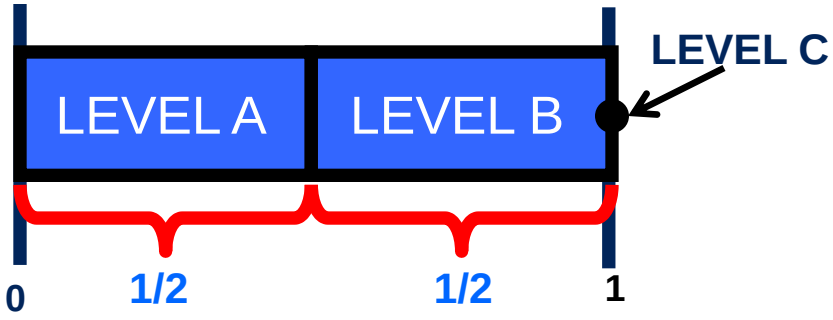
Correspond to the ratio of $p(C_1)$'s for each level

Example: Modification of Probabilities with More Information

Suppose data (i.e., more information) establish that the Level C condition has $p(C_1) = 0.95$

Table F-2: 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	0.2375
B	Partial clarity	3	0.7125
C	Unclear	4	0.95



$S = 1/2 \log_2 [2] + 1/2 \log_2 [1/.475] + 0 \log_2 [1/0] = 1.0$

Additional information modifies state table probabilities

CENTROIDS OF INTERVALS AT 0.25, 0.75, and 1 or a ratio of 1:3:4

Correspond to the ratio of $p(C_1)$'s for each level

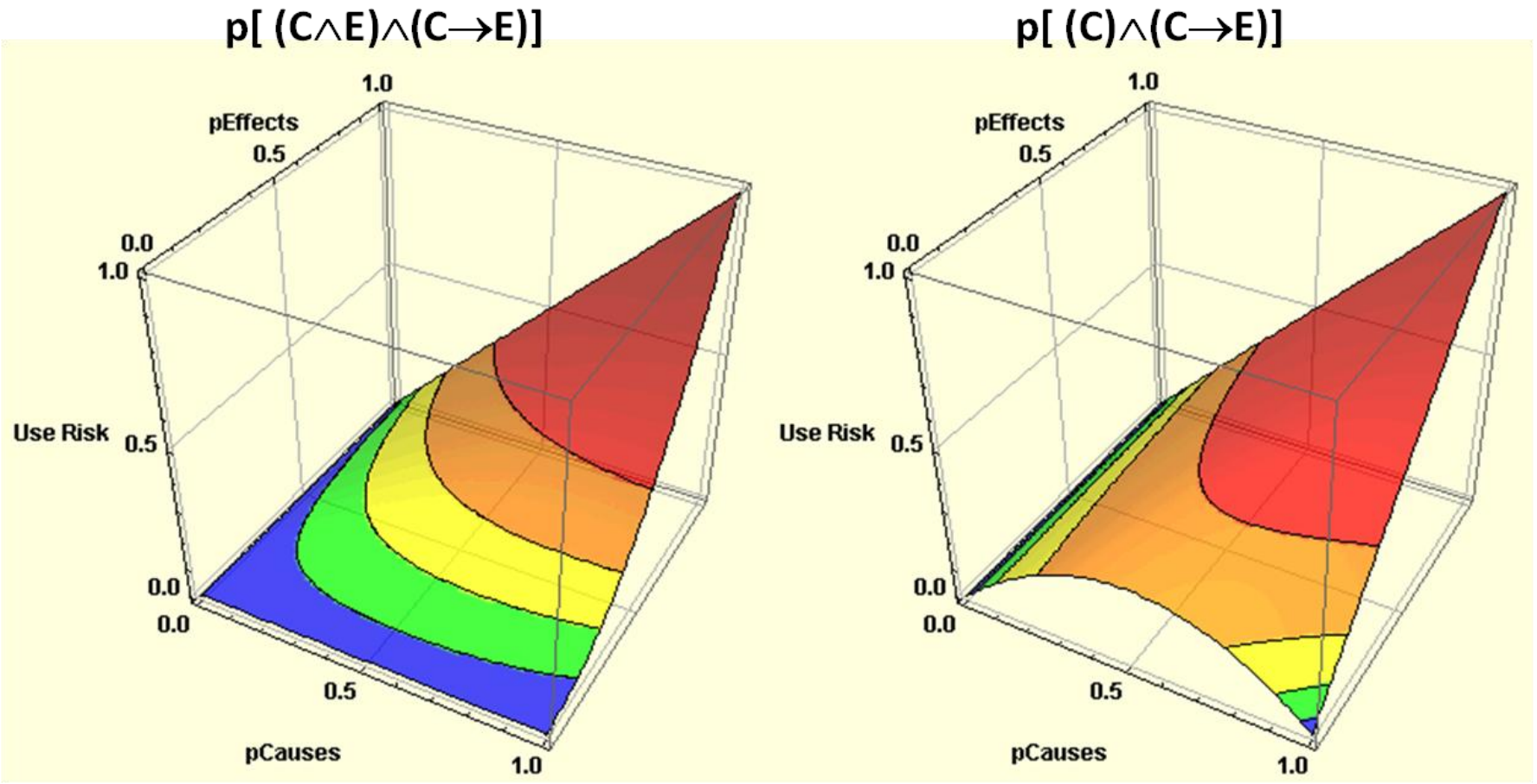
Comments about MURM

- MURM allows full use of *ALL* available information in Use Risk assessment: objective (quantitative), limited (such as only knowing ranking of alternatives), and subjective
- Use of information entropy theory reduces or precludes inadvertent & intended bias in assessment
- MURM operates at M&S individual capability levels (vice only treating M&S results as a whole)
- MURM uses the V&V Composite Model (VCM) to identify V&V phases, activities, and tasks comprehensively
- MURM uses the V&V Process Maturity Model (VPMM) characterization of information quality & objectivity to relate information produced by V&V activities to UR

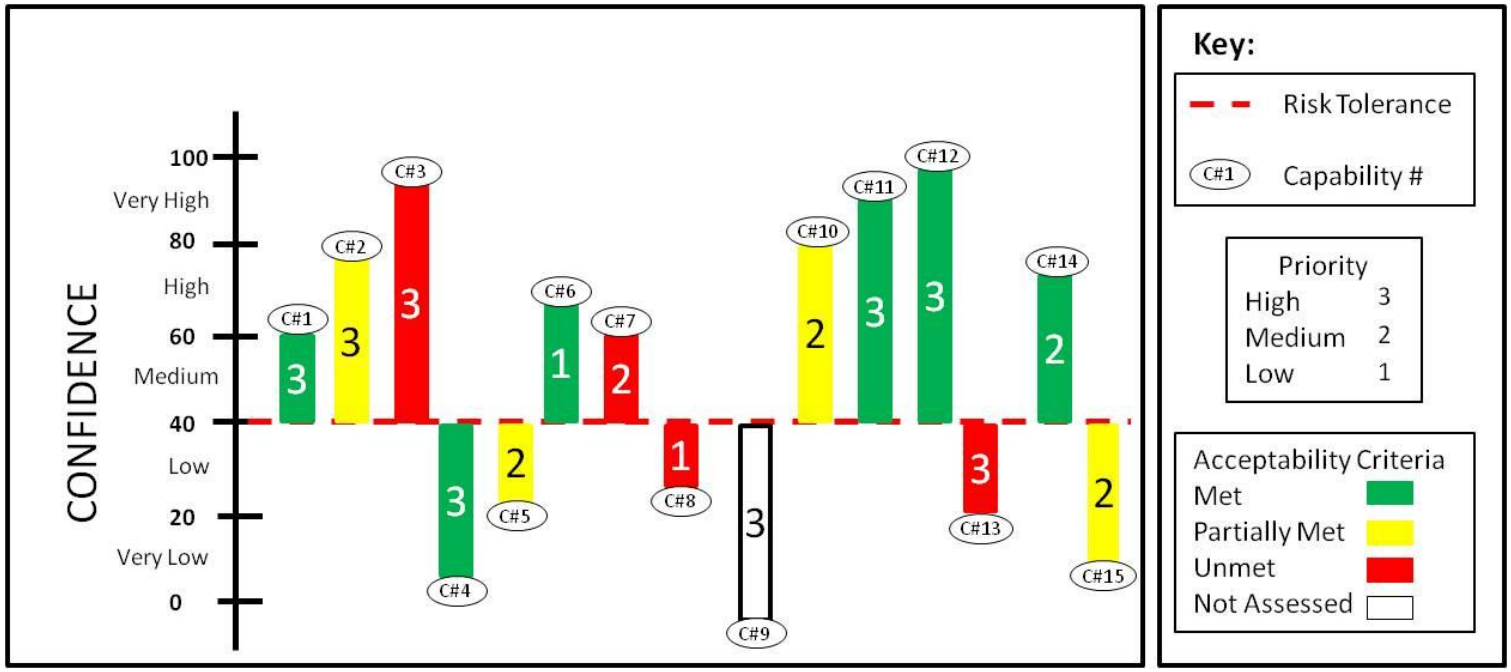
Comments about MURM (cont.)

- A V&V Techniques Catalogue characterizes the quality and objectivity of information that can be produced by the various V&V techniques
- State tables support Use Risk computation – can be developed to the level supported by available information
- Any sophisticated math methodology requires particular math skills. With automation, V&V personnel will be able to use MURM just as *Mathematica* is used effectively by those without the math skills to apply all of its techniques without use of that program.
- Examples of the kinds of information that can be produced by MURM are illustrated on following charts.

Two M&S Use Risk Definitions Compared Illustrating Descriptive Power of MURM



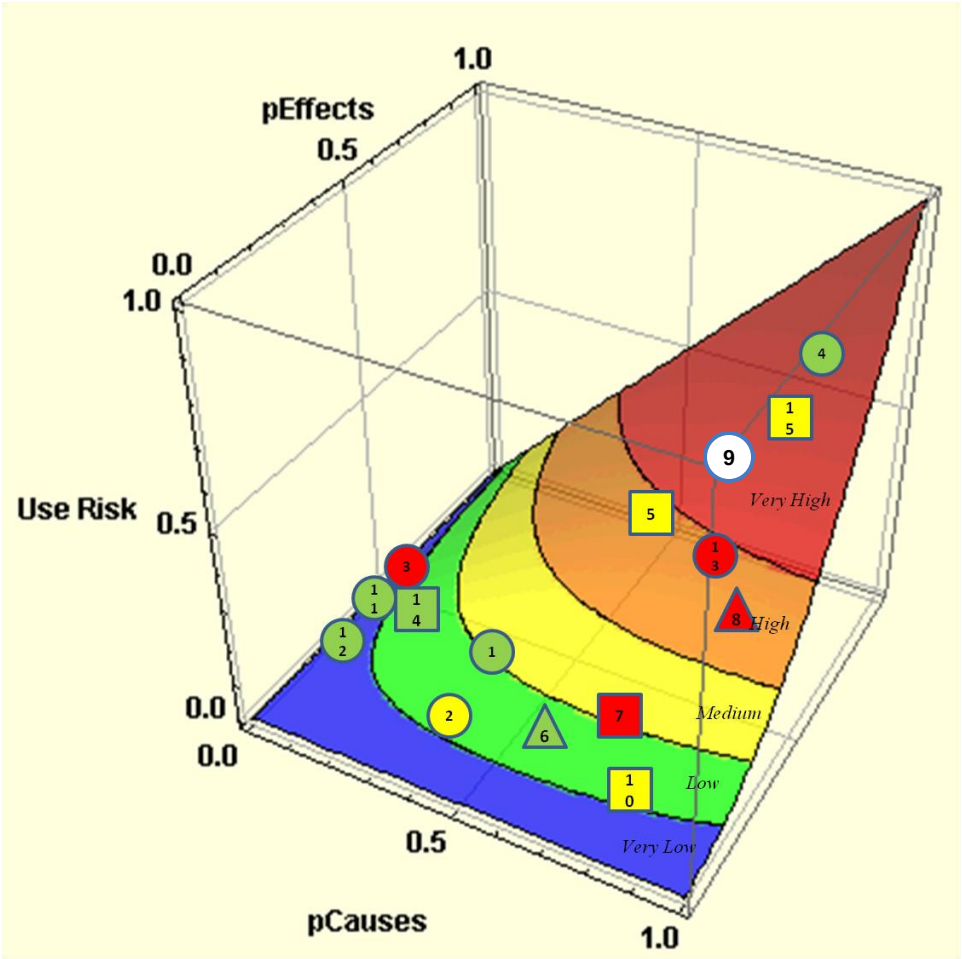
Visualization of MURM C_3 Results (see following chart for full MURM results)



This figure 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 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.

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.)

VISUALIZATION OF MURM RESULTS



Illustrated are results for 15 capabilities & their associated M&S User Risk

A capability is identified by a number within its symbol

A capability's priority is noted by the shape of its symbol

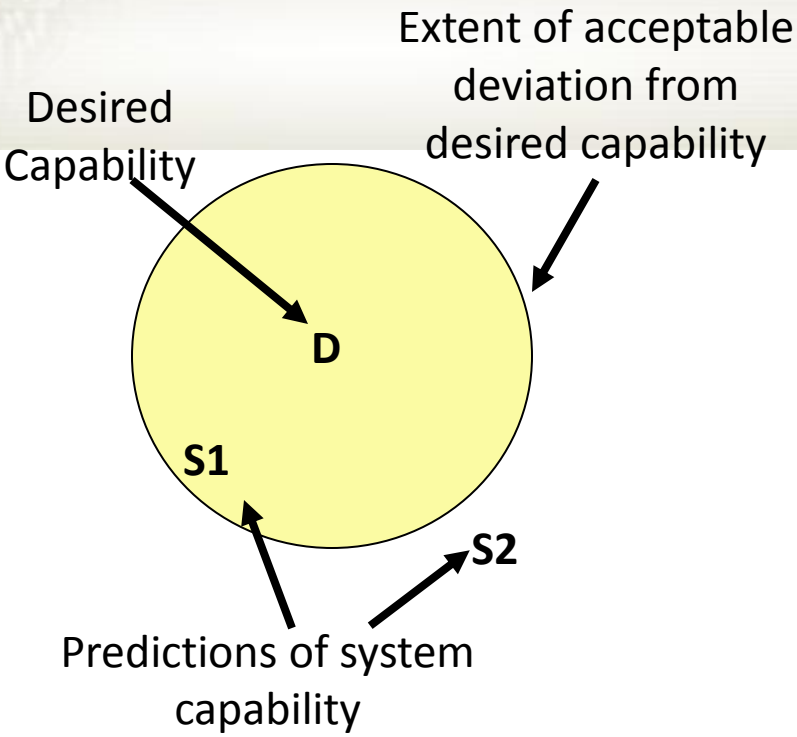
A capability's disposition is noted by the color of its fill

The position of the capability's symbol on the risk surface indicates its M&S User Risk; the surface is divided into five equal area regions and are labeled from Very Low risk to Very High risk.

The distribution of the symbols give a sense of the User Risk status of the M&S; for example, #4, a high priority capability, met its acceptability criteria but the risk level, perhaps driven in part by the $p(C_3)$ recommendation confidence, propels it into the high user risk region.

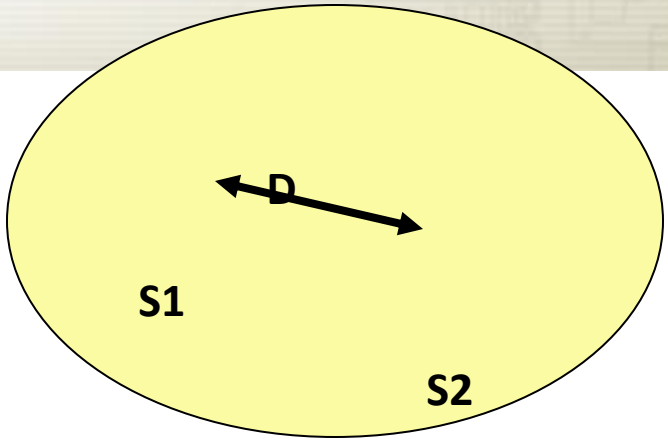
Some Thoughts on Uncertainty

Conclusions Change with Uncertainties Considered



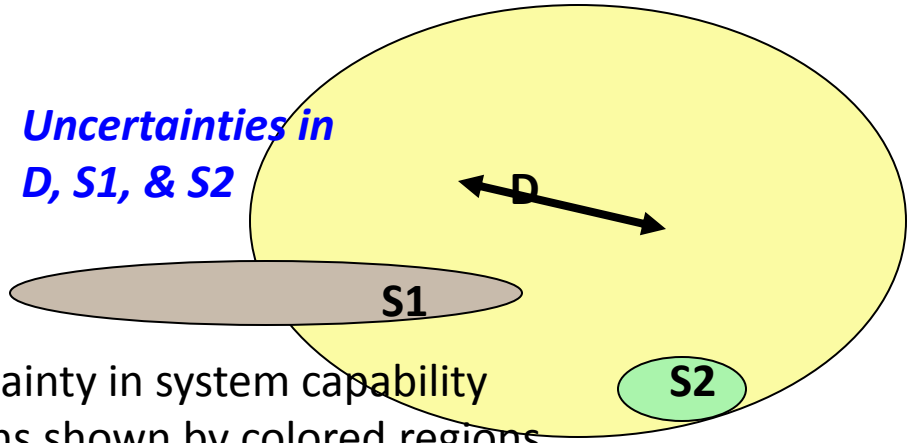
No Uncertainties

Uncertainty in D location
(indicated by double arrow)



Region of acceptable deviation based upon uncertainty in location of D

Uncertainties in D, S1, & S2

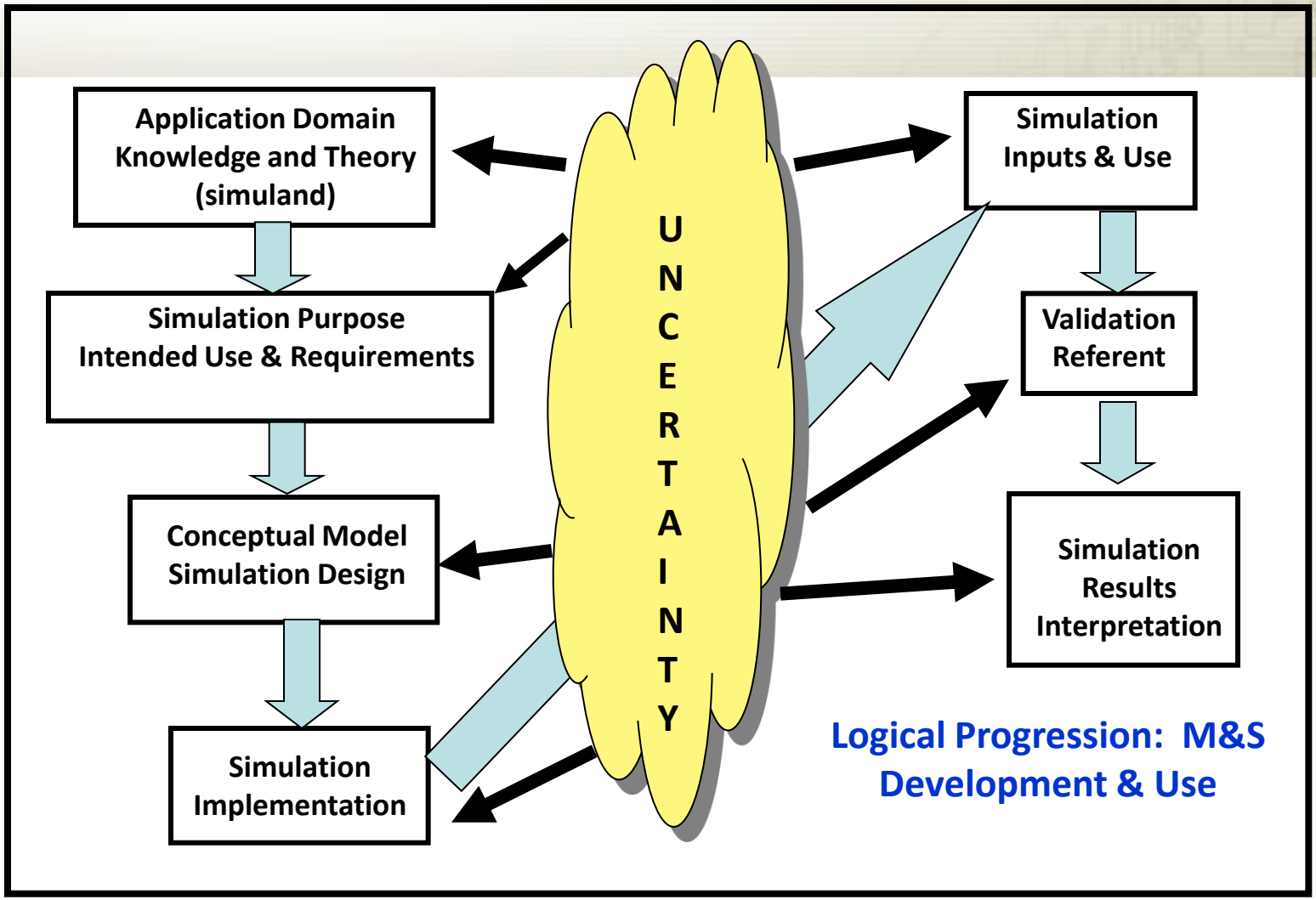


Uncertainty in system capability predictions shown by colored regions about S1 and S2

Perspective on M&S Uncertainty

- Lots of good work is being done in UQ (uncertainty quantification), BUT that work has focused on identification and quantification of uncertainty re particular parameters with consideration of how it is propagated through M&S – not on comprehensive treatment of M&S uncertainty
- In 2009 a paradigm to facilitate comprehensive consideration of M&S uncertainty was developed in conjunction with work for DTRA related to assessment of medical resource M&S related to WMD situations
- When M&S uncertainty is addressed comprehensively, uncertainty in M&S results can be large, even orders of magnitude in some application domains

2009 SIW Paper M&S Uncertainty Paradigm (09S-SIW-082)



Objective: Comprehensive Perspective on M&S Uncertainty

Observations

Some aspects of M&S uncertainty are generally ignored or left unspecified in various M&S communities, such as:

- User effects (same code, same problem, different users => different results) – Nuclear Power Plant community is the only group that gives this kind of uncertainty much attention; others ignore it
- Judgment uncertainty (from SME use, filtering of test data, etc.) is very seldom characterized or quantified
- Uncertainty impact of where M&S application space is sampled for the validation referent on use of the M&S for other areas in its application space is seldom addressed explicitly – this kind of consideration has many dimensions, such as scalability and validity of M&S for extrapolation beyond regions of test results

Use of a comprehensive paradigm, such as the one shown, can help to ensure that all aspects of M&S uncertainty are addressed and not simply ignored or neglected

Incomplete consideration of M&S uncertainty increases M&S use risk

More Information Is Available

For detailed discussion of MURM and its application or to obtain the MURM report:

Contact Simone Youngblood (JHU/APL) at

Simone.Youngblood@jhuapl.edu
(240) 228-7958