A Common M&S Credibility Criteria-set Supports Multiple Problem Domains

Joe Hale
NASA
Marshal Space Flight Center
Huntsville, AL
256-544-2193
joe.hale@nasa.gov

Bobby Hartway
AEgis Technologies Group
631 Discovery Drive
Huntsville, AL 35806
256-922-0802
b.hartwar@aegistg.com

Danny Thomas
AEgis Technologies Group
631 Discovery Drive
Huntsville, AL 35806
256-922-0802
danny.thomas@aegistg.com

Keywords:
"Modeling and Simulation Tools", "NASA", "Integrated Management", "Taxonomy", "Credibility Creteria"

ABSTRACT: *Credibility management of M&S depends on two factors. They are how well the analysts know the credibility of the M&S and how clearly that knowledge is presented to decision makers. An easily understood measure of credibility is needed. Conversely, aerospace simulations are necessarily very complex and difficult to characterize. We present a hierarchical system of credibility measurement that can suit the most straightforward assessment requirements for top management while retaining the details to provide useful feedback to developers. At the top level, the scale is a single number. This "one-dimensional" scale is useful for top management. The second level is defined using a systems engineering, multi-dimensional view of the conceptual solution space whose domains encompass the important attributes of the assessment. The solution space that results has three orthogonal dimensions:*

1) *The Problem Domain, which addresses how well the M&S fit the intended use;*
2) *The M&S Development Domain, which addresses how well the M&S is built with respect to the intended use; and*
3) *The User / Analyst Domain, which addresses how well the M&S used with respect to the intended use.*

Subsequent indentures divide these domains into a few hundred generic criteria that are selectively applied by subject matter experts (SMEs) to assess the credibility status of M&S. These criteria are tailored (mapped) one time for applicability to each subject domain of interest, and to the domain agency's standard(s) for quality (or confidence or credibility, etc. as required). This tailored set is then further tailored for each M&S assessment in that domain. Only criteria that apply to any given M&S application are used.This paper presenst applications of this "common/tailored criteria concept" to examples for NASA space exploration and MDA missile defense.

1. INTRODUCTION

The primary goal of NASA-STD (I) 7009, Standard for Models and Simulations, is to ensure that credibility of results from models and simulations (M&S) are properly reported to those making critical decisions. The secondary goal of the standard is to assure that the credibility of the results from M&S meets project requirements. The standard addresses development and application of M&S, as well as analysis, documentation, and presentation of the results from M&S. It applies to M&S used for Design and Analysis; Natural Phenomena Prediction; and Manufacturing, Assembly, Test, Operations and Evaluation.

Credibility of M&S may be defined using a systems engineering, multi-dimensional view of the conceptual solution space whose domains encompass the important attributes of the assessment. The solution space that results has three orthogonal dimensions as shown in Figure 1:

1) The Problem Domain, which addresses how well the M&S fit the intended use (Intended Use, or IU)
2) The M&S Development Domain, which addresses how well the M&S is built (Built Well, or BW) with respect to the intended use; and
3) The User / Analyst Domain, which addresses how well the M&S is used (Used Right, or UR) with respect to the intended use.

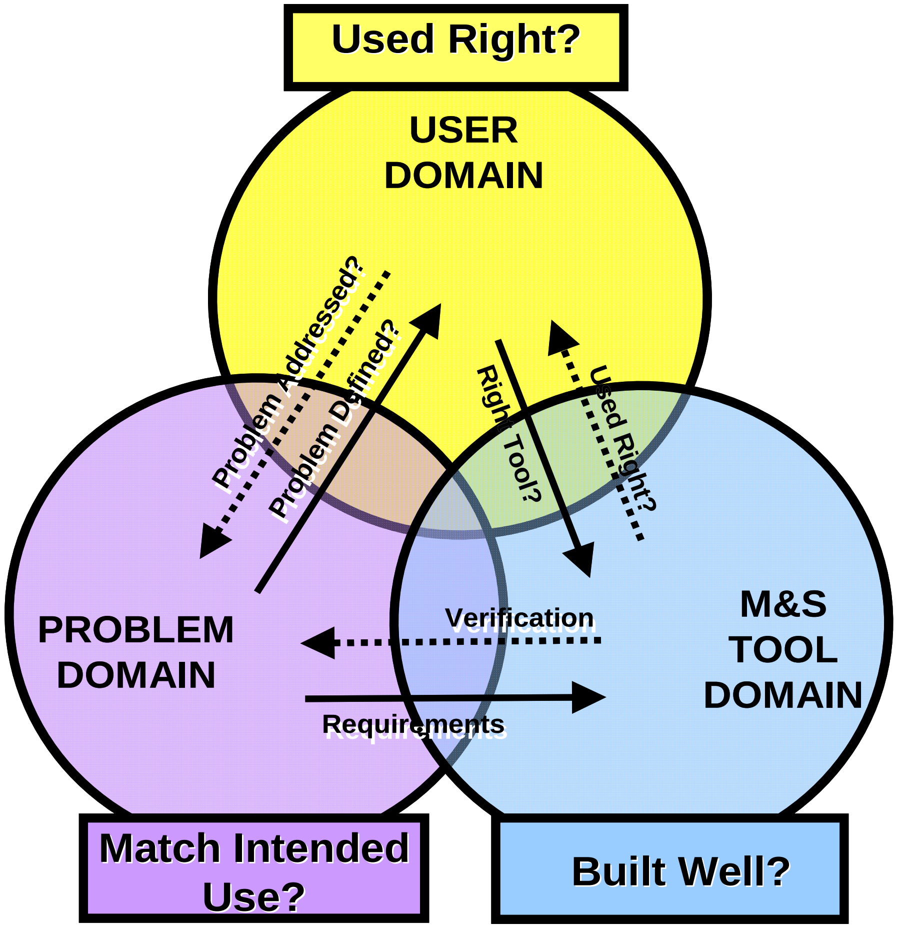

Figure 1. The Three "Domains" of M&S Tool Evaluation-Criteria

2. Assessing Credibility Against Requirements

2.1 Acceptable Achievement

Defined credibility criteria measure the simulation in absolute terms of goodness, answering the question, "How good is it?" These measures are important to simulation developers who want to improve their products and managers who want to improve the overall quality of groups of simulations. Decision makers who use simulation results, however, seldom have perfect simulations to use, so their question often becomes, "Is it good enough?". They can accept less than perfect achievement in some measures in favor of stressing measures that are critically important to their current intended use. Figure 2.1 depicts the relationship of these two important questions. Assessing simulation credibility requires a combination of a top-down approach to set credibility requirements and a bottom-up approach to measure the credibility of the simulation against these requirements.

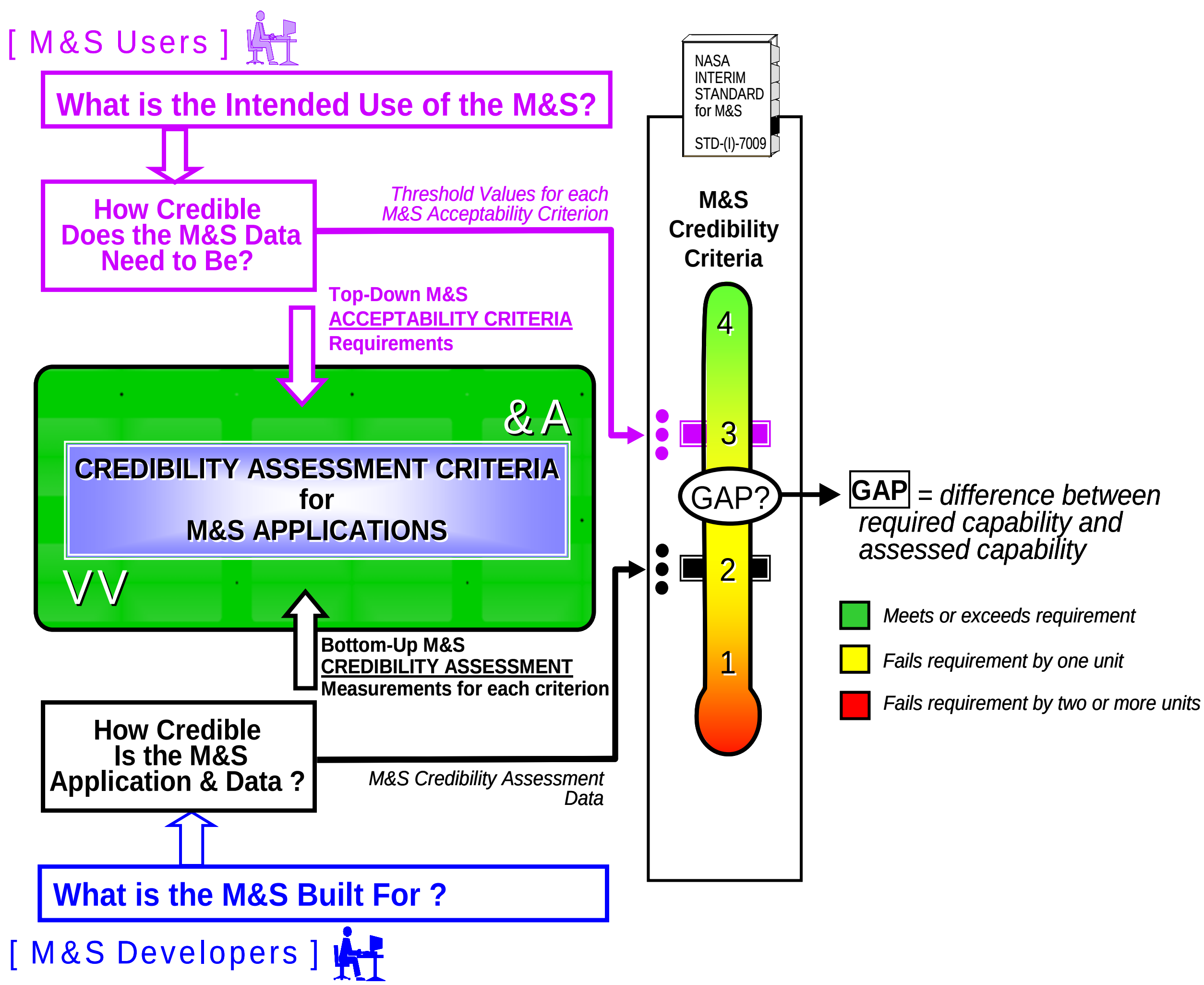

Figure 2. 1 Assessing Simulations against a Specific Need

3. Establishing Thresholds

The Accreditation Authority establishes a set of credibility criteria "thresholds" that represents the minimum acceptable credibility needed in each evaluation area for this specific simulation, at this specific time, and for this specific use. Figure 3 shows the minimum acceptable credibility thresholds as the horizontal value bars for each sub-criterion.

Typically, the decision maker will want the highest confidence in credibility for every category of a simulation. Realistically, time and resources will not permit full attention on all criteria; trades will need to be done to prioritize where to allocate efforts. Therefore, the decision maker should prioritize the criteria in order of importance. In this example, sub-criteria "Right Entities" and "Right Functions" in the Fit Intended Use Domain, sub-criterion "Validated Outputs" in the Built Well Domain, and sub-criterion "Operator/Analyst Qualified" in the Used Right Domain are set to the highest level.

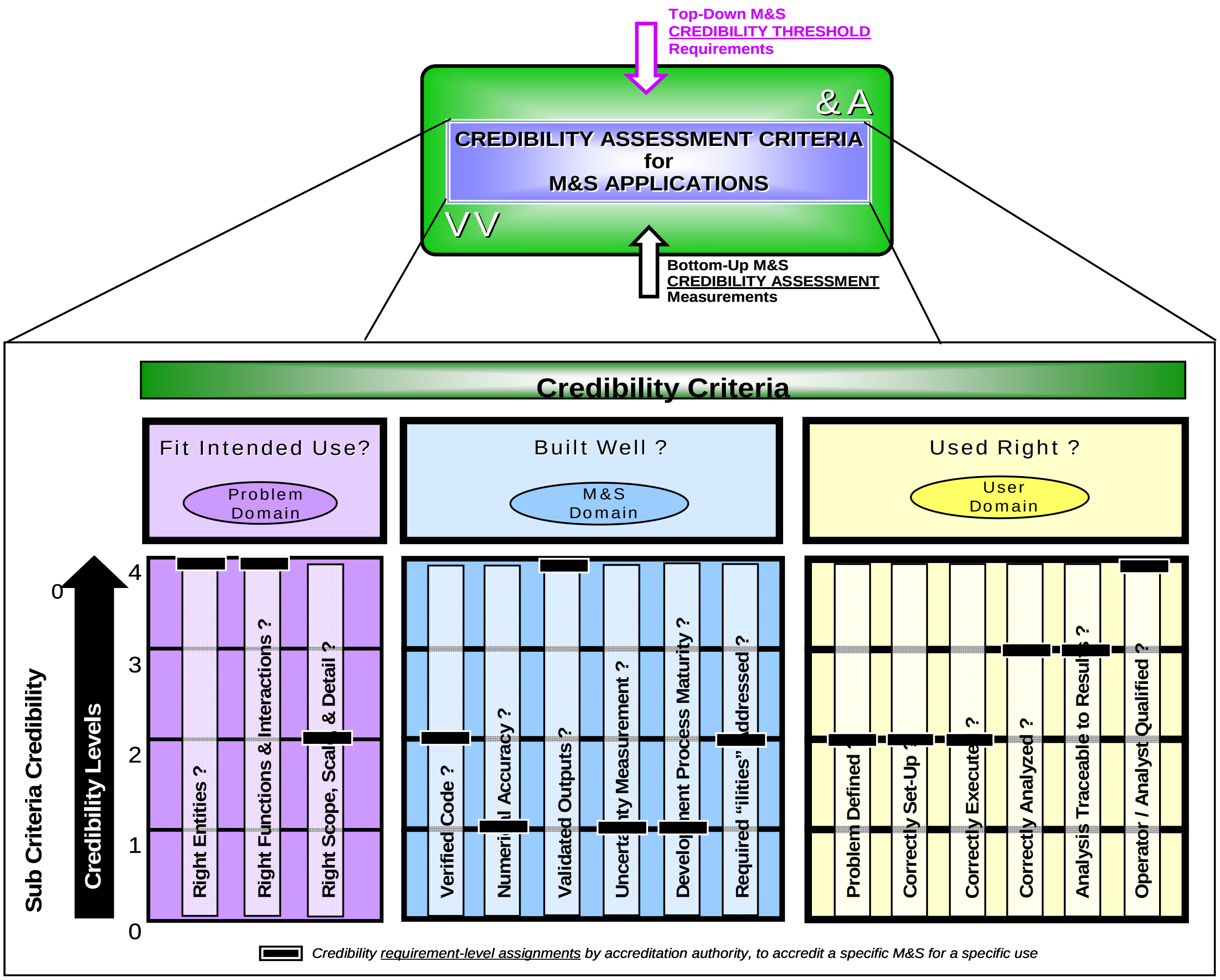

Figure 3. Credibility Thresholds Applied to Simulation Criteria

4. Assessing Achievement

Analysts assess each criterion by well-established methods. Each criterion has a credibility level from one to four. Returning to the example in Figure 4, the "Right Entities" and "Validated Outputs" subcriteria scored very low relative to the required credibility thresholds .

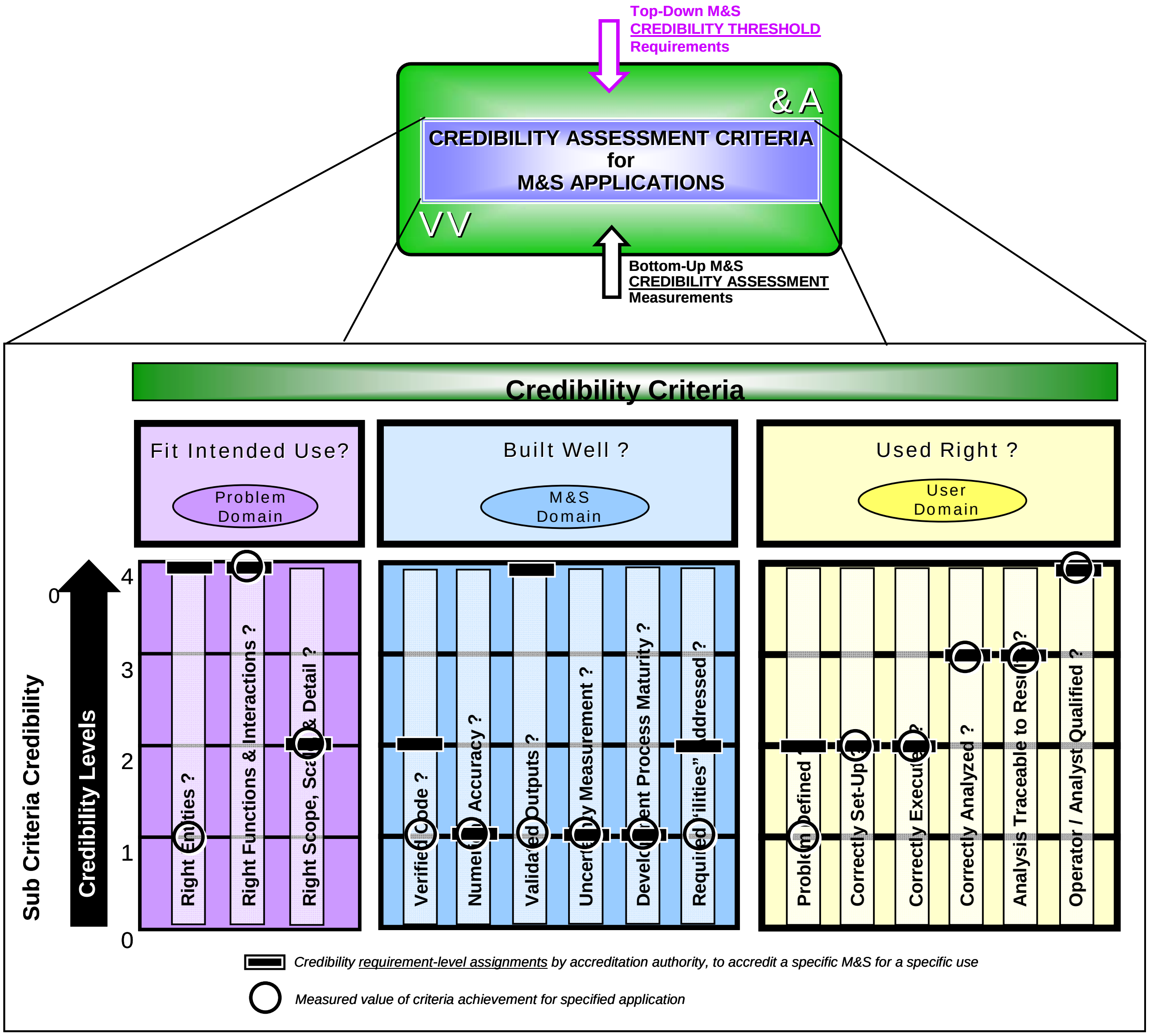

Figure 4. Example Credibility-Achievement Assessment

5. Evaluating Results

A simple subtraction of the achieved credibility levels from the required credibility levels yields a "gap" in credibility level between what is desired and what is measured as shown in Figure 5-1 Credibility Gap Analysis.

It is important to note that the gap may arise from an actual shortfall in the simulation credibility, or it may stem from incomplete measurement. If the achieved level meets or exceeds the required level, no gap exists. If the achieved falls one level below that required, a gap of one and a "yellow" condition exists. If the achieved is two or three levels below that required, gaps of two and three exist and the condition is "red." This quickly shows the problem areas on the example. There is a worst-case gap of three (3) in Fit Intended Use, a worst-case gap of three (3) in Built Well, and a worst-case gap of one (1) in Used Right.

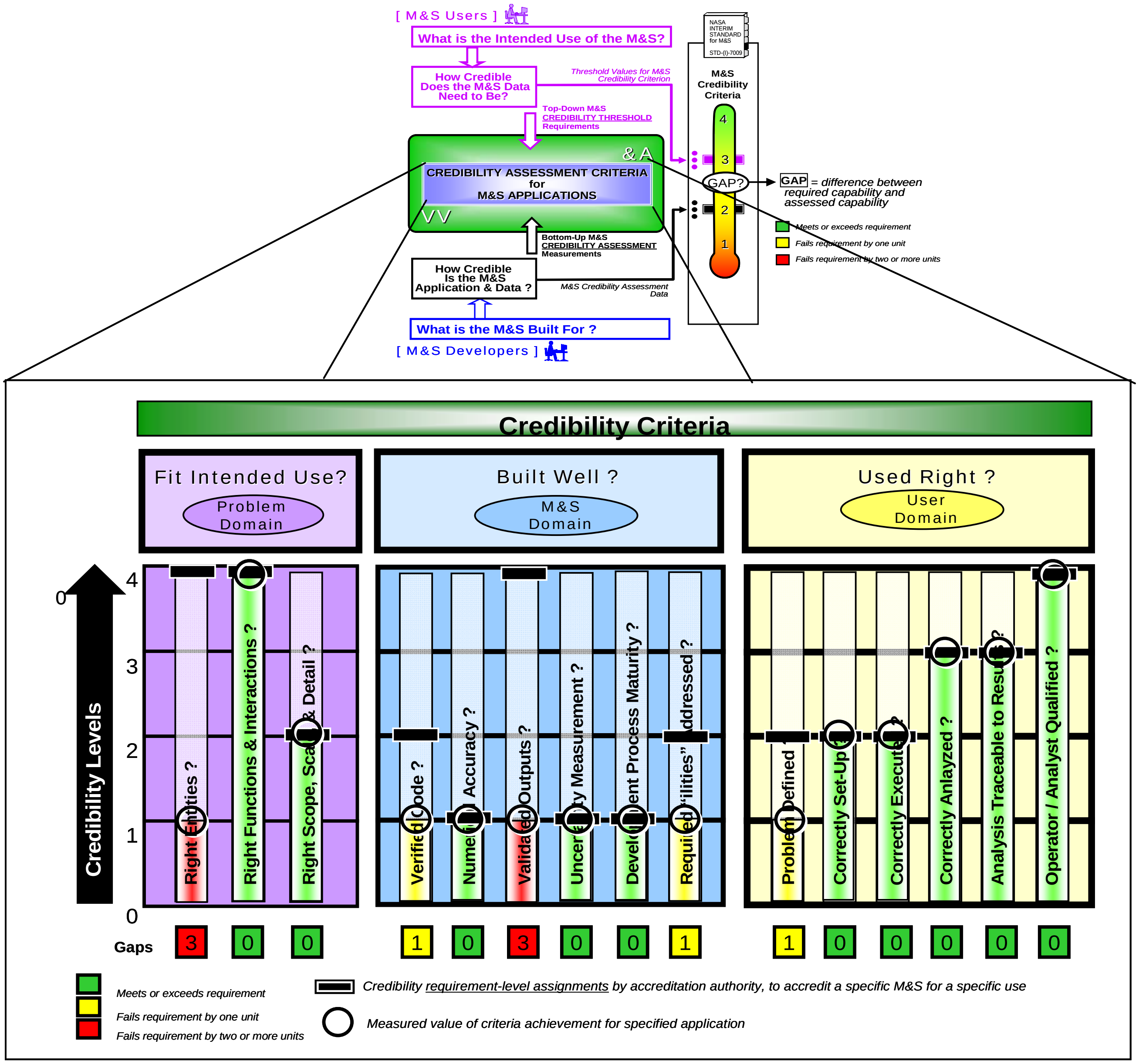

Figure 5-1. Credibility Gap Analysis

One determines the credibility level for each of the three domains by converting to the 1 to 4 credibility-level scale by subtracting the worst-case gap from a perfect credibility level score of 4. Hence, in this example the credibility level scores for the three domains are 1, 1, and 3 (see Figure 5-2).

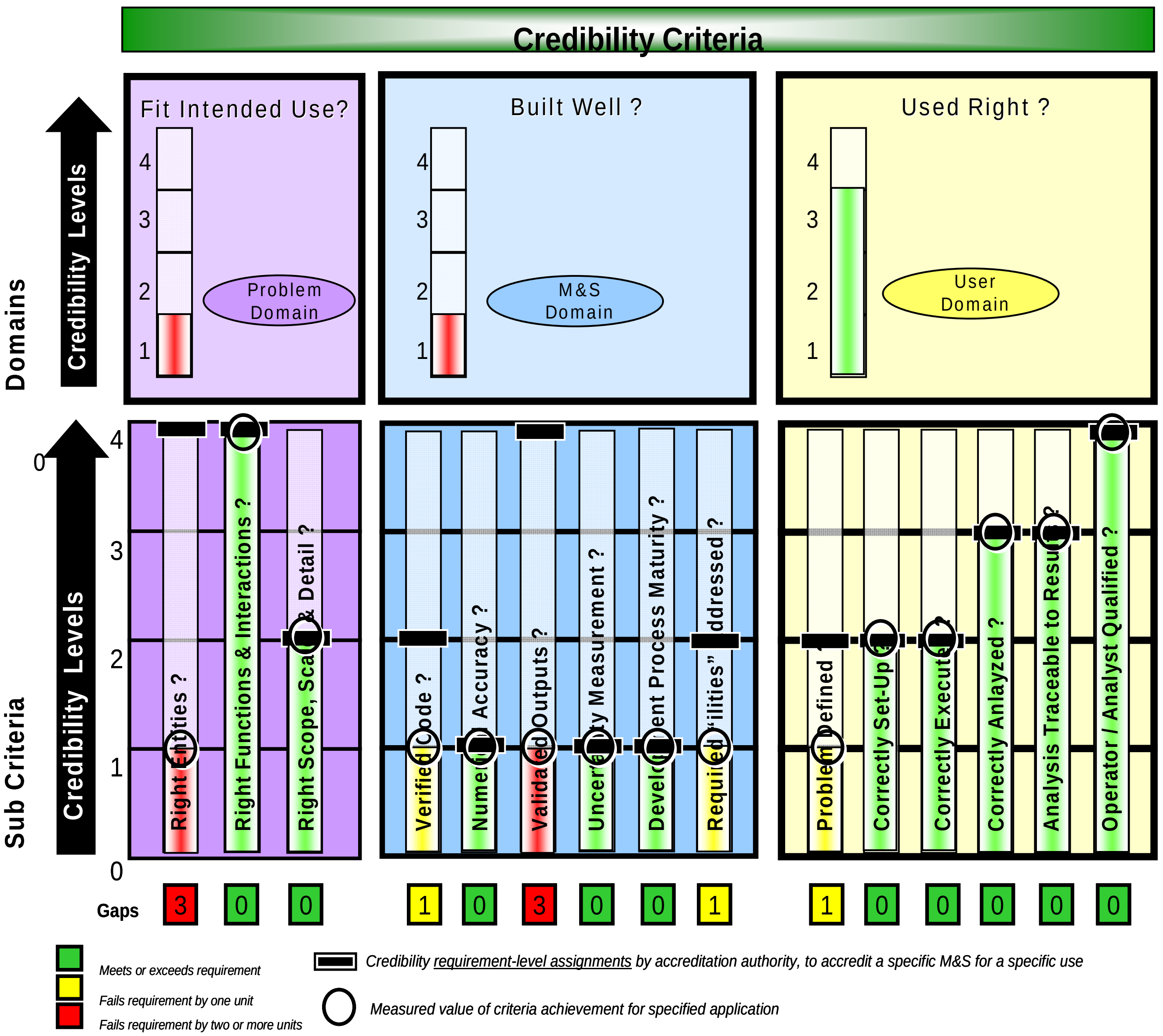

Figure 5-2. Credibility Gap Analysis at Domain Level

Finally, the lowest credibility level score among the three Domains becomes the Credibility Level reported on the one-dimensional scale. Thus, the one-dimensional Credibility Level in this example is one (see Figure 5-3).

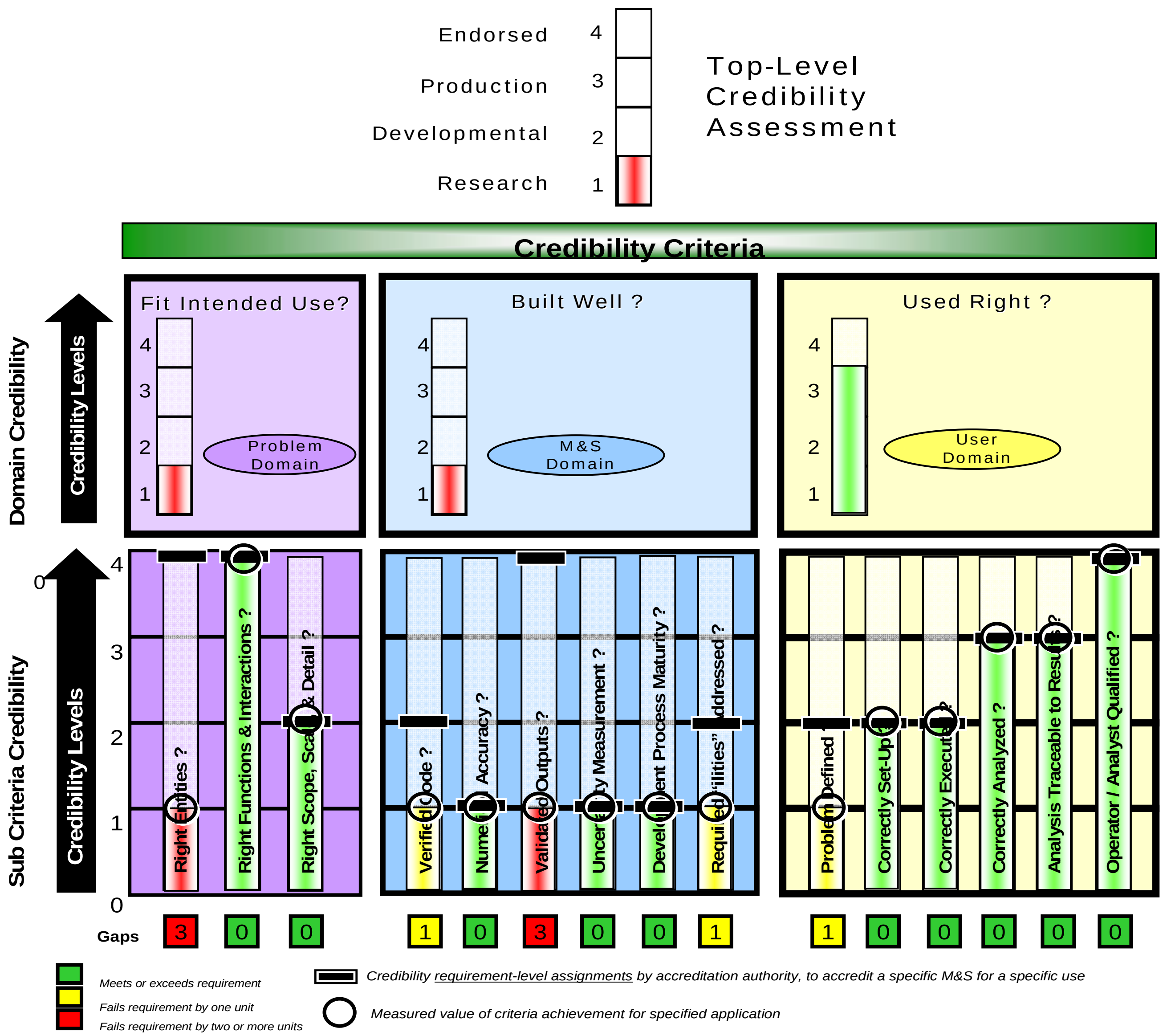

Figure 5-3. Credibility Gap Analysis at One-Dimensional Level

6. References

[1] NASA-STD (I) 7009, Standard for Models and Simulations

Author Biographies

JOSEPH PETER HALE, II has worked with NASA at the Marshall Space Flight Center (MSFC) since 1985. He received a B.A. in Psychology from the University of Virginia in 1976 and an M.S. in Applied Behavior Science (Psychology) from Virginia Tech in 1981. He continued graduate studies at Virginia Tech through 1984 completing all coursework for a Ph.D. in Human Factors Engineering while participating as a Graduate Co-op Student at MSFC. In 1990, he received an M.S. in Systems Management from the Florida Institute of Technology.

Mr. Hale is currently Lead for the Exploration Systems Mission Directorate's Integrated Modeling and Simulation Verification, Validation, and Accreditation activity. Prior to that he was the Lead Systems Engineer for the Next Generation Launch Technologies' Advanced Engineering Environment. Much of his prior work at MSFC was as a Human Factors Engineer, working various projects, including Spacelab and the International Space Station. He spent five years as Team Lead for the Human Engineering and Analysis Team.

Mr. Hale is a Certified Human Factors Professional (CHFP) (Board of Certification in Professional Ergonomics), is a founding member and first president of the Tennessee Valley Chapter of the Human Factors and Ergonomics Society (HFES).

BOBBY HARTWAY is a Senior Research Scientist with AEgis Technologies Group in Huntsville, Alabama. He has developed a new paradigm for simulation characterization and requirements development for space and defense systems. He is using this paradigm to support NASA's activities for integrated management of modeling and simulation. Mr Hartway is a Certified Modelind and Simulation Professional.

DANNY THOMAS is a Senior Research Scientist with AEgis Technologies Group in Huntsville, Alabama. He is supporting NASA's effort to institute consistent management practices for simulation development and use. He has developed simulations for space and defense. Mr Thomas is a Certified Modelind and Simulation Professional.