Generic Methodology for Verification and Validation (GM-VV) – Part 1: Introduction and Concepts

Dr. Manfred Roza
National Aerospace Laboratory NLR
Q-tility V&V Expertise Center
Anthony Fokkerweg 2, Amsterdam
NETHERLANDS

Manfred.roza@nlr.nl or Manfred.roza@q-tility.nl

ABSTRACT

The Generic Methodology for Verification and Validation (GM-VV) is a generic and comprehensive methodology for structuring, organizing and managing the Verification and Validation (V&V) of M&S assets. The GM-VV is a new recommended practice within the Simulation Interoperability Standards Organization (SISO), which is the result of a joint development effort with NATO. The GM-VV provides a technical framework to efficiently develop arguments to justify why M&S assets are acceptable or unacceptable for a specific intended use. This argumentation supports M&S stakeholders in their acceptance decision-making process regarding the development, application and reuse of such M&S assets. The GM-VV technical framework assures that during the execution of the V&V work the decisions, actions, information and evidence underlying such acceptance arguments will be traceable, reproducible, transparent and documented.

This paper is the first in a series of three papers that collectively describe and illustrate the complete GM-VV technical framework and its application. This first paper presents the rationale for the GM-VV and its development, an overview of the technical framework, and the basics concepts part of this framework.

1.0 RATIONALE FOR GM-VV

Increasingly, models and simulations (M&S) are developed and deployed as enabling technology to support system analysis, design, test and evaluation, acquisition, training and education. Not only M&S is used more and more frequently for such purposes, but also its users rely more and more on M&S technology or solutions as the primary or even solely means to address their specific problems in this regard. Therefore, within safety critical domains it is imperative to perform systematic and robust verification and validation (V&V) to ensure that both the development and utilization of M&S technology is cost-effective, and provide credible results that do not pose unacceptable risks. For example, in civil flight crew training this is reflected by the flight simulator qualification standards and regulations that are imposed by various national aviation authorities.

However, experience outside this specific area shows that V&V is often more of an afterthought than an integral part of any M&S development, employment and procurement policy. This is due to the fact that V&V for M&S is still a relatively new field of technology and practice, with many very divergent approaches. The V&V method that works best in a given situation depends on the individual needs and constraints of an M&S organization, project, application domain or technology. Therefore, there exist many different approaches to V&V that rely on a wide variety of V&V terms, concepts, products, processes, tools or techniques. In many cases the resulting proliferation restricts or even impedes the transition of V&V assets and results from one M&S organization, project, and technology or application domain to another. In particular, in the military domain this is a large issue due to the increasing usage of distributed simulation technology such as DIS, HLA or TENA. These technologies facilitate the reuse and interoperation of M&S

systems and M&S system components from various suppliers, application domains and nations, within a single simulation exercise to replicate joint, combined or allied operations.

This context was the key driver behind the development of the Generic Methodology for Verification and Validation (GM-VV) [1],[2],[3] within the Simulation Interoperability Standards Organization (SISO) in cooperation with NATO.

2.0 DEVELOPMENT OF GM-VV

The development of the GM-VV started in 2003 within the two Western European Armament Group (WEAG) projects called Reference for Verification, Validation and Acceptance (REVVA-1 and REVVA-2) (Figure 1). In this group Defence R&D organizations from Canada, Denmark, Germany, France, Italy(REVVA1), The Netherlands and Sweden joined their R&D capabilities to study and develop a common reference base for performing verification, validation and acceptance of models, simulations and data. It was this group that developed the underlying concepts and technical framework for the GM-VV. Its final products REVVA recommended practice guide (RPG), REVVA Handbook and REVVA reference manual were finalized and delivered in 2008 by a dedicated tiger-team of VVA experts [10],[11]. During REVVA-2 the products became mature enough to start a standardization process within SISO such that the GM-VV would become available to a larger M&S community. This resulted in the establishment of the GM-VV Product Development Group (PDG) within SISO that transformed the REVVA products into a set of standard guidance products in accordance with the SISO policies and procedures. To support and further mature the GM-VV within SISO, the REVVA-2 project was succeeded by the task group 073 from NATO Modeling and Simulation Group (MSG-073) [12] in which Germany and Turkey joined the team The major contributions of the NMSG-073 is the execution of series of (inter)national case-studies to test the GM-VV in real M&S applications and further refine the GM-VV with the lessons-learned from these studies.

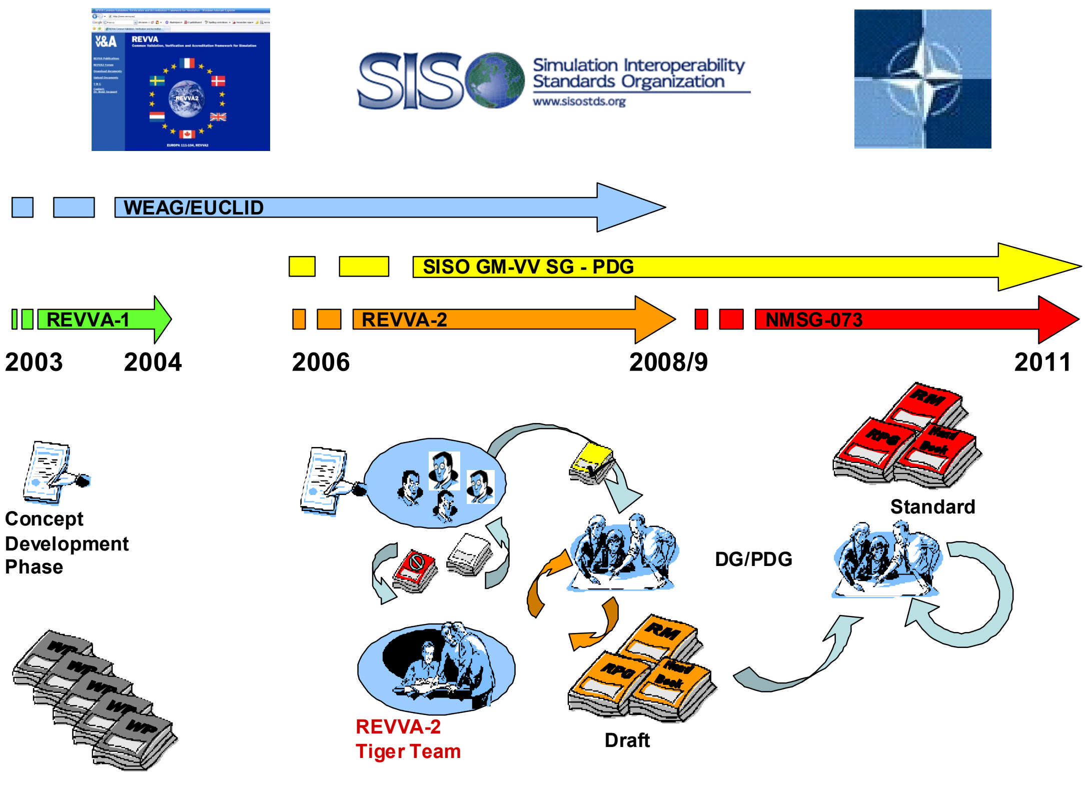

Figure 1: GM-VV Development Life-Cycle.

Therefore, it can be concluded that the GM-VV is the result of a long international cooperation effort that started with a European effort evolving to a cooperative effort of NATO and SISO, supported by many nations, M&S organizations and industry. The GM-VV is also the first completed joint development of SISO and NATO since the signature of a technical cooperation agreement between these organizations in 2007.

The GM-VV development life-cycle (Figure 1) was guided by the following starting-points:

- The GM-VV shall provide common semantics and components for V&V that can be used unambiguously across and between different M&S organizations, projects, technology or application domains.
- The methodology scope shall encompass the V&V aspects of all artefacts within the M&S life-cycle, ranging from real world needs, through M&S development and usage, to supporting the acceptance decision.
- The methodology shall be suited for performing V&V activities concurrently to the M&S development process, as well as for post-hoc V&V of existing M&S assets, and it shall be applicable to a wide variety of M&S technologies and application domains.
- The methodology shall provide support for establishing V&V agreements between V&V user and supplier organizations. Such a V&V agreement covers all requirements and other arrangements placed on a V&V effort.
- The methodology shall be applicable not only on a technical level but also on a project and enterprise level to address managerial and organizational aspects of V&V efforts.
- The methodology shall be V&V client oriented and product-driven. For each V&V product an activity shall be defined to produce that product and for each activity a role shall be defined.
- The methodology shall be able to address various levels of organizational independence depending on the V&V client needs.
- The methodology shall facilitate the development of traceable, transparent and reproducible evidence-based arguments that underpin an acceptance recommendation.
- The methodology shall provide tailorable V&V products, activities and roles with respect to V&V needs, use risks and available resources to obtain a cost-effective V&V effort.

3.0 GM-VV DOCUMENT SET

The GM-VV development life-cycle resulted in three SISO products (Figure 2); two guidance products that were balloted and one reference product that was carefully reviewed. This GM-VV document set was completed and officially approved by SISO in October 2013.

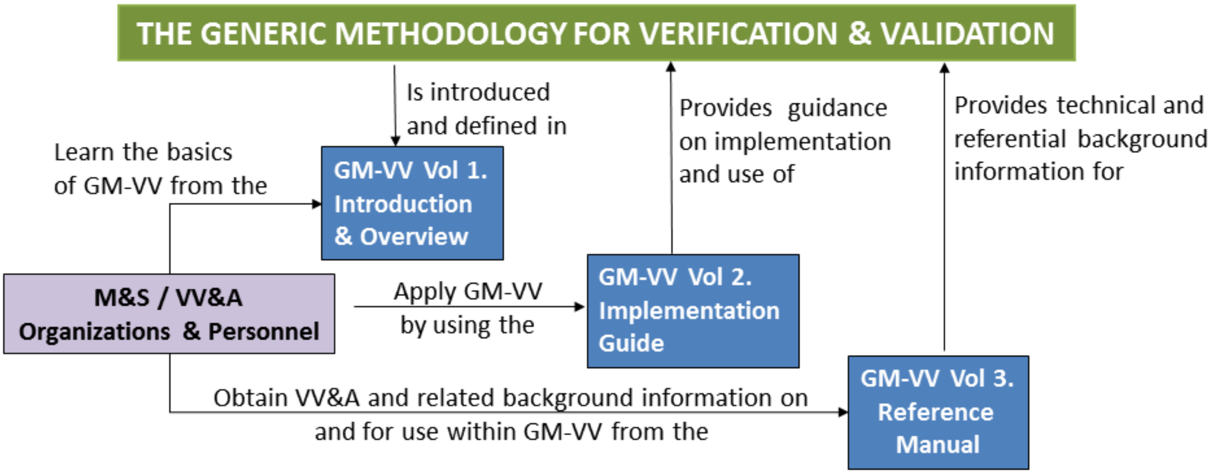

Figure 2: GM-VV Document Set.

GM-VV volume one gives a high-level overview to the GM-VV technical framework (Chapter 4) and its underlying concepts [1]. This document is intended for all M&S professionals, managers and users/sponsors, who are trying to gain knowledge about the importance and benefits of incorporating V&V within their organization or projects, and seeking a general applicable and standardized V&V methodology. This document is highly recommended for newcomers to the topic of V&V of M&S that want to learn and understand the basic V&V terminology, concepts and application principles before immersing themselves in the specific details of V&V applications. GM-VV volume one provides its readers the basics of the GM-VV and introduces them to the two other GM-VV documents.

GM-VV volume two describes the technical core of the methodology [2]. This document contains details on the implementation framework components, as well as detailed guidance on how to apply these components in conjunction with the tailoring framework to develop concrete V&V solutions. GM-VV volume two aims at M&S professionals that are responsible for embedding V&V practices and standards in their organizations, managing and executing V&V activities as part of their M&S projects.

GM-VV volume three is not a guidance product but a reference product [3], which aims to document the foundations of the GM-VV concepts, their dependencies and rationale. It provides references to technical literature for the methodology itself and V&V methods and techniques that can be used in conjunction with the GM-VV. Most important, this reference manual provides the practical case-studies as conducted by the MSG-073 members and provide hands-on experiences and lessons-learned on how to best apply the GM-VV with M&S projects and organizations.

4.0 GM-VV TECHNICAL FRAMEWORK OVERVIEW

The GM-VV aims not to replace any existing V&V approaches, methodologies, standards or policies of M&S organizations, technology and application domains; nor it is intended to be prescriptive, in that it does not specify a single concrete or unique solution for all V&V applications. Instead, the purpose of the GM-VV is to provide a general baseline and guidance for VV&A of M&S that:

- Facilitates a common understanding, communication, comparison and interoperability of native V&V practices and standards.
- Is applicable and tailorable towards individual V&V needs of a wide variety of M&S organizations, project, technologies and application domains.

To accommodate the above mentioned objectives and comply with the starting-points mentioned in Chapter 2, a reference model and architecture approach was used for the design of the methodology. This approach was applied in such a way that the GM-VV is not directly tied to any specific M&S application domain, standard, technology, organization or other distinctive M&S implementation details for V&V. As a result, the GM-VV comprises an abstract technical framework that consists of three parts that build upon existing V&V methods [4],[5],[6],[7] and other related practices [7] (Figure 3). The conceptual framework provides unifying terminology, concepts and principles to facilitate communication, common understanding and execution of V&V within an M&S context (Chapter 5). The implementation framework translates these concepts into a set of generic architectural template and building blocks for the development of concrete and consistent V&V solutions supporting an individual M&S organization, project, and technology or application domain. GM-VV provides a tailoring framework that utilizes these building blocks to develop and cost-efficiently apply such V&V application instantiations, based on M&S project or organization requirements for V&V (e.g., acceptance needs, budget, time frame, risks or available resources). As such, the GM-VV provides a high-level framework for developing concrete V&V solutions and conducting V&V into which lower-level practices (e.g., tools, techniques, tasks, acceptability criteria, documentation templates) native to each individual M&S organization, project, technology or application domain can easily be integrated.

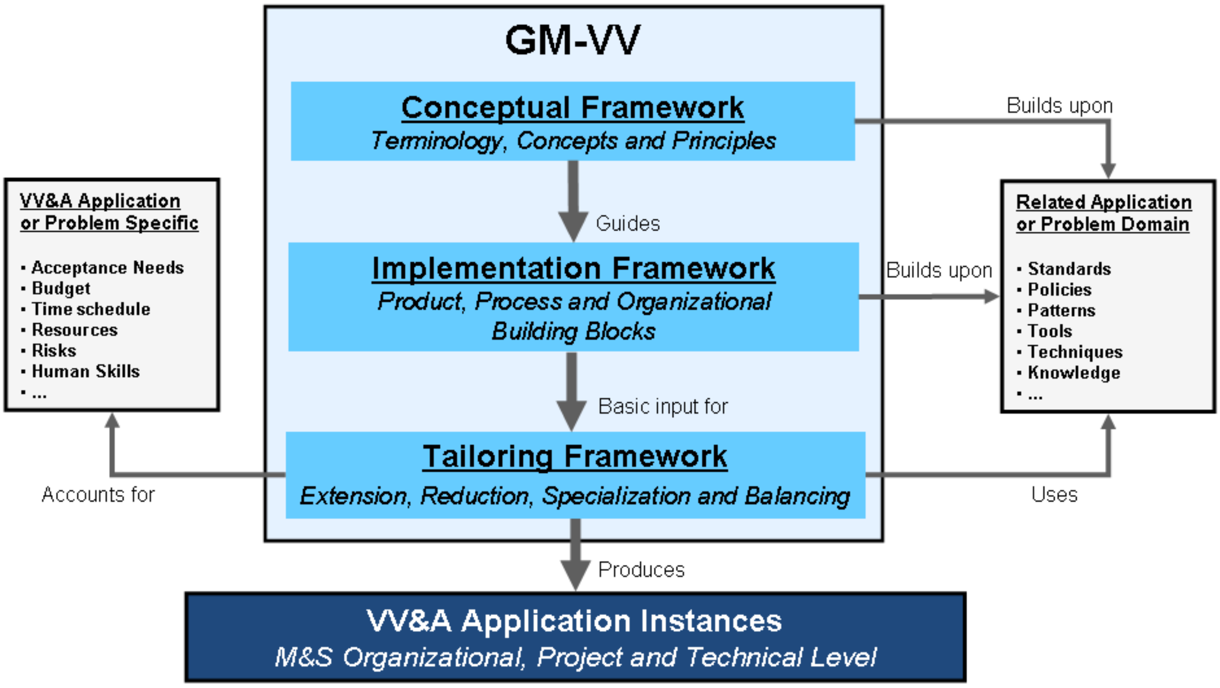

Figure 3: GM-VV Technical Framework Design and Operational Usage Concept.

The GM-VV implementation framework and tailoring framework are discussed in detail within the two other accompanying papers on GM-VV [13],[14]. The remainder of this paper will focus on the GM-VV conceptual framework.

5.0 GM-VV CONCEPTUAL FRAMEWORK

The GM-VV conceptual framework provides fundamental and general applicable terminology, semantics, concepts and principles for V&V. The framework aims to facilitate communication, understanding and implementation of V&V across and between different M&S contexts. The conceptual framework is the fundament upon which the GM-VV implementation framework rests (Figure 3). In this chapter the basic concepts of this framework will be introduced and discussed in detail.

5.1 M&S and its Linkage to System Engineering

M&S are closely related to systems engineering. A possible definition of a model is that it is an abstract representation or specification of a system. Abstraction in this context is a process in which a relative sparse set of relevant (sub)systems, relationships and their inherent qualities are extracted or separated from a more complex reality. A model can represent a system that exists in our material world but also non-existing or virtual systems or hybrid combinations thereof. That part of (the imagined) reality that the model is supposed to represent is called the simuland.

In a simulation the model is used to replicate the simuland behavior. Thus a simulation is a method, software framework or system to implement and evaluate a model over time i.e., it is a system in which a model is made to execute and is exercised. This model in its executable form is called the M&S system. Within the GM-VV, M&S systems are therefore considered to be systems of systems that have a lifecycle and are subject to system engineering practices. Moreover, models and simulations are considered to be part of a larger system in which they are used. From this perspective, M&S is a systems engineering specialization.

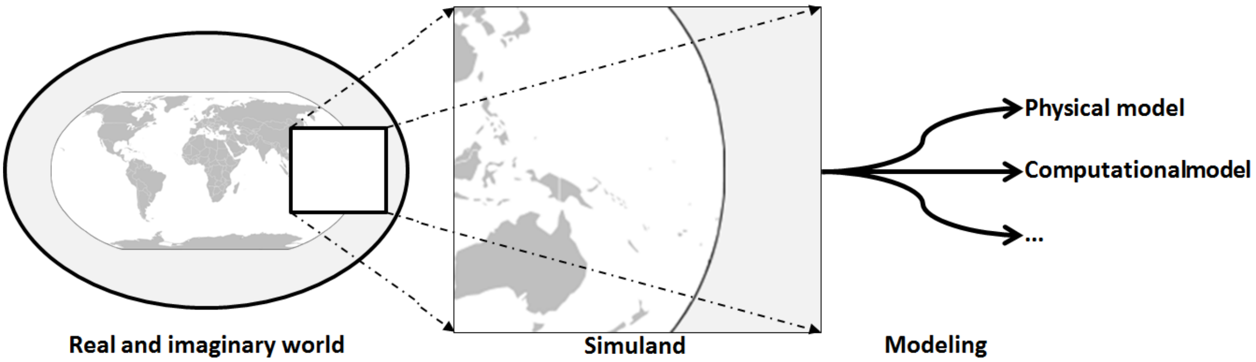

Figure 4: That part of the real and imagined world that is being modeled is called the simuland.

The M&S system is provided with input and its output is used within a certain context provided by a frame system, the frame is called the *Simulation Frame*. The model that is executed in the simulation is controlled and observed by means of its ports. Through these ports simulation data, stimuli or settings are entered into the model and simulation output leaving the executed model is observed. During the simulation the model behaves according to a dynamics that represent the state change and behavioral properties of the simuland. The notion of time, behavioral representation and frame are fundamental characteristics of a simulation [3].

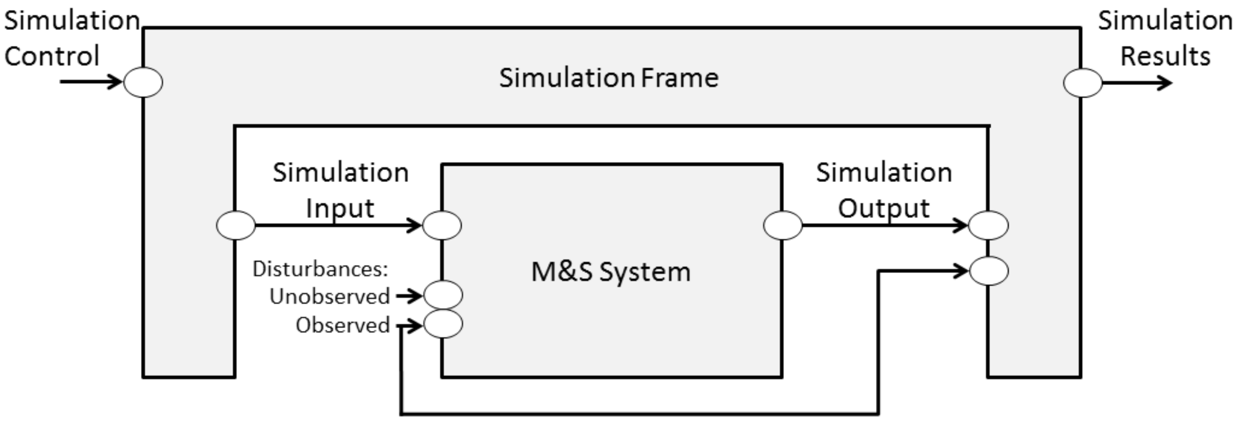

Figure 5: Relation Between Simulation Frame and the M&S System.

To properly replicate the simuland for the intended use, the model is configured, controlled and stimulated by the Simulation Frame by means of input trajectories, scenario's, parameters, environment variable settings and experimental control settings. Furthermore, environment disturbances may impact the behavior of the M&S system. During the execution of the model, human input can become part of the displayed behavior. This can be from trainees, but also from operators such as opponent players to provide stimuli to the trainees or Subject Matter Experts (SMEs) that interfere with the execution of the simulation for some purpose dictated by the Simulation Frame (e.g., keeping the execution within a desired regime).

V&V is an intrinsic part of the systems engineering process. Therefore, the GM-VV considers the V&V of M&S as a specialization of systems engineering V&V. There is however, a difference between general systems engineering and systems engineering as applied to M&S. In M&S, abstractions of the simuland are used with the aim to replicate the simuland behavior. The additional work for V&V of M&S is to provide evidence to show that these abstractions produce behavior that replaces the simuland behavior to such an extent that the M&S system or its results fit its intended use (i.e., validity). Hence, the GM-VV can be integrated with, complement or extend the V&V processes within such existing systems engineering methodologies or standards.

5.2 M&S Uncertainty, Risks, Confidence, Credibility and Trustworthiness

In systems engineering the interim products resulting from the life cycle steps may be faulty for various reasons. This is true for general systems engineering, but even more so for M&S because of the modeling process. M&S is considered as being more an art than a science. Errors are either introduced from standard systems engineering process, the modeling process and/or errors in the employment of the M&S System can lead to uncertainties about the utility of M&S. These uncertainties ultimately lead to (operational) risks in the real world e.g. inadequate trained aircraft maintenance person that making mistakes because of being trained with an invalid M&S system. The relevant concepts are briefly discussed as follows.

5.2.1 Uncertainty

Uncertainty is an inclusive term. It covers the lack of certainty, whether the uncertainty can be modeled probabilistically or not. This definition allows the term uncertainty to be applied to anything. Different communities restrict the application of this term to, for example, predictions of future events, physical measurements already made, and unknowns. Two major types of uncertainty can be defined: uncertainty due to the lack of knowledge (i.e., epistemic uncertainty) and uncertainty due to non-deterministic behavior (i.e., stochastic uncertainty).

5.2.2 Risk

Risk is a concept that denotes the probability of specific undesired eventualities. There are many definitions of risk that vary by specific application and situational context. However in general usage the convention is to focus only on potential, negative impact to some aspect of value that may arise from a future event. In engineering, the definition of risk is a factor that could result in future negative consequences: usually expressed in terms of impact and likelihood or probability. Risk is thus a state of uncertainty where some possible occurrence may cause an event that results in a loss, catastrophe, or other undesirable outcome. Risk in the GM-VV context can be divided into two types: project risks and use risks. Project risks are risks directly related to the project time and budget, while user risks are directly related to the use of M&S system for a specific intended purpose.

5.2.3 Confidence, Credibility, Trustworthiness

Confidence is generally described as the state of being certain, that either a hypothesis or prediction is correct, or that a chosen course of action is the best or most effective. Credibility in general refers to the trustworthiness of a source or a message. A survey on the terms credibility, confidence and trustworthiness shows that they are often used interchangeably within the M&S domain.

All aforementioned definitions have several implications in common that complicate the objective assignment of any measure of credibility to a model or simulation. For the M&S domain credibility is nearly synonymous with confidence and trustworthiness. Credibility is a property of the information being presented that involves the belief of the observer or recipient of that information. Therefore, the perception of credibility is inherently subjective and cannot be meaningfully measured. Furthermore, credibility is only loosely coupled to the process for deriving the presented information. Therefore, the integrity of that process can only contribute to credibility if the observer understands that process and appreciates its limitations. Finally, in order for the observer to trust the credibility of the process for producing information they must also trust that the people who applied that process did so correctly.

Within the GM-VV the terms credibility, confidence and trustworthiness are also used interchangeably and considered to have the same semantics.

5.3 Landscaping the V&V World

The basic premise of GM-VV is that models and simulations are always developed and employed to help fulfill the specific needs of their stakeholders (e.g. trainers, analysts, and decision makers). GM-VV uses a four-world view to structure this larger context (Figure 2). Together, these four worlds define a generic M&S life-cycle and process view. When this process is properly followed the M&S-based solution that is transferred to the real world for operational usage should fulfill the original needs.

Within the four world context, stakeholders (e.g. developers, managers and authorities) exist who are responsible for making acceptance decisions on the use of M&S systems, its results or any intermediate products. Such stakeholders are referred in the GM-VV as the *V&V User/Sponsor*. The inherent problem is that it is not possible to demonstrate with absolute certainty that such M&S assets will meet the needs *prior* to its actual use. Consequently, there is a probability that the M&S-based solution will not meet user needs and hence, it poses a risk. An M&S asset is *acceptable* when the responsible stakeholder has sufficient confidence in the success of the asset without posing unacceptable risks (e.g. costs, liabilities). The objective of verification and validation, as defined by GM-VV, is to collect, generate, and maintain a body of evidence. This accumulated evidence is then applied to build an argument to support acceptance decision-making processes. *Validation* in GM-VV is referred to as the process to ensure that the right M&S asset is built or procured for the intended use (i.e. M&S validity). To ensure that the M&S asset at delivery can be demonstrated to be valid, it is necessary that the M&S asset is built and employed in the right manner.

Verification is therefore referred to in GM-VV as the process to ensure that the evolving M&S asset is built right (i.e. M&S correctness).

V&V is a specific M&S problem domain that is known within GM-VV as the V&V World (Figure 6). The V&V world groups the products, processes and organizational aspects necessary to develop a suitable acceptance recommendation for the responsible stakeholder in one of the other four worlds. This recommendation included in a *V&V Report* is the key deliverable of a V&V effort and contains evidence-based arguments regarding the acceptability of an M&S system or result. The V&V effort is driven by the *V&V Needs* that are traceable to the stakeholders acceptance decision needs (e.g. budget, responsibilities, risks, liabilities).

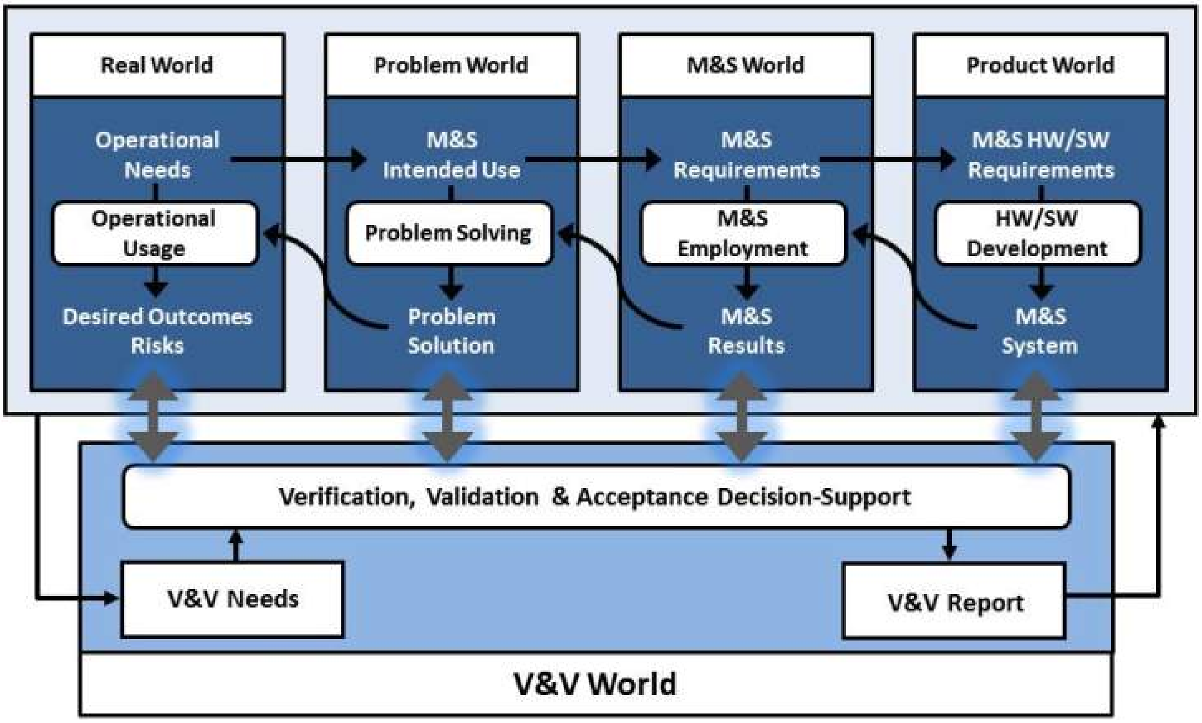

Figure 6: GM-VV Worlds View.

The development of an acceptance recommendation in the V&V world is driven by the V&V needs that are traceable to the V&V User/Sponsor's acceptance decision or procedure(s) needs (e.g., budget, responsibilities, risks, liabilities). Therefore, the extent, rigor and timeframe of a V&V effort depend on these needs. Depending on these needs, the V&V effort could span the whole or specific M&S lifecycle phase of the four worlds; could focus on one specific or multiple (intermediate) M&S products; and should match the development paradigm that was used (e.g., waterfall, spiral). Each case may require a separate acceptance recommendation with its own scope and development timeline. Moreover, the way the V&V effort interacts with the four M&S-based problem worlds also varies from case to case. These mutual dependencies are depicted in Figure 6 with bidirectional arrows that interface the V&V world with each of the four M&S-based problem solving worlds. Two classical types of V&V that can be identified based on the time frame of their execution are:

- *Post-hoc V&V*: V&V conducted in retrospect on an M&S system after development or on M&S results after M&S system employment.
- *Concurrent V&V*: V&V conducted in prospective throughout the whole M&S lifecycle to manage and improve the quality of newly developed M&S systems or results.

The GM-VV supports both V&V time frames but is not limited to these distinct types. A V&V effort can be post-hoc, concurrent, iterative, recursive or even be a recurrent effort in the case where legacy M&S products are updated or reused for a different intended-use.

5.4 Acceptability Criteria Satisfaction and Evidential Quality

The V&V objective is to develop an acceptance recommendation that convincingly shows why an M&S asset is acceptable or not acceptable for the stakeholder. This objective is articulated in GM-VV as an *acceptance goal*. This goal is conceptually accomplished in GM-VV by means of five high-level activities. First, define a set of concrete and assessable *acceptability criteria* for the M&S asset. Second, collect or generate relevant evidence to demonstrate the satisfaction of the acceptability criteria. Third, assess the evidential quality of this demonstration. Fourth, based on the outcomes of the previous three steps develop arguments underlying claims whether or not the M&S asset is acceptable for its intended use (i.e. *acceptance claim*). Finally, compile all previous information into an acceptance recommendation for the stakeholder.

As depicted in Figure 7, GM-VV defines three major M&S asset property areas for which acceptability criteria can be defined. Here *utility* refers to the usefulness of the M&S asset in solving the needs from the real world. *Utility* properties could comprise sub-types such as M&S asset value, cost and use risks. *Validity* properties refer to the degree of realism (i.e. fidelity) of the system of interest that is replicated by the M&S asset [9]. The validity significantly determines the utility of an M&S asset. *Correctness* properties refer to the extent to which the M&S asset implementation and usage conforms to its specifications (e.g. conceptual model, design); and is free of design, development and employment defects (e.g. semantic errors, syntactic errors, numerical errors, user errors). M&S asset correctness impacts its utility and validity.

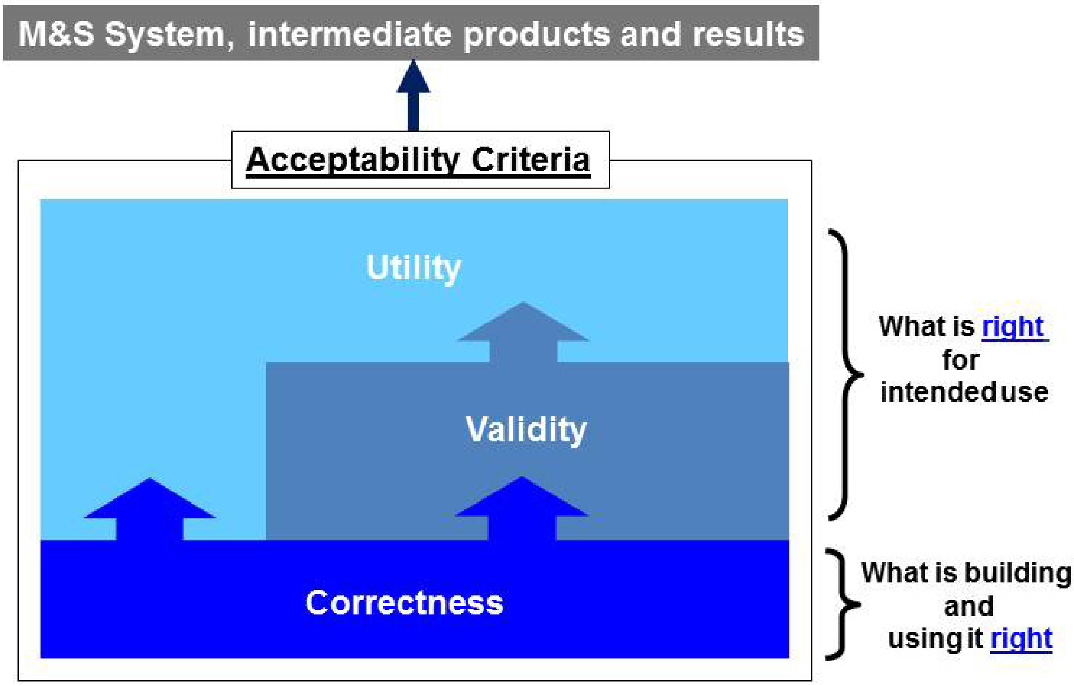

Figure 7: Acceptability Criteria.

To be able to make a well informed acceptance decision, also the evidential strength of an acceptance claim must be known. The latter determines the level of trust that can be placed on such claim, and directly relates to the stakeholders tolerance of use risk. For this purpose, the GM-VV identifies V&V quality properties that refer to how well the V&V effort is performed (e.g. rigor) with regard to developing the acceptability criteria, collecting evidence, and assessing to what extent the M&S assets satisfy the acceptability criteria. Examples of V&V quality properties are the completeness, consistency and relevance of the acceptability criteria. In the process of collecting and assessing evidence, V&V quality properties may include: knowledge and data uncertainties, skill level of V&V personnel, reliability and repeatability of V&V techniques, relevance and justification for any assumption made in this whole effort.

The eventual acceptance recommendation comprises acceptance claims along with the supporting arguments and underlying evidence. An acceptance recommendation is not necessarily just a yes or no claim. Meeting all the acceptability criteria means the claim can be made that the M&S asset should be accepted for the intended use without limitations. If all acceptability criteria are not met, alternative weaker acceptance claims with supporting arguments and evidence can be constructed. Such alternative acceptance claims could, for example, provide recommendations regarding conditions or restrictions under which the M&S asset can still be used; or on modifications that, when implemented, will lead to an unconditional acceptance for the intended use.

5.5 V&V Argumentation Approach

Evidence and arguments underlying an acceptance recommendation should be developed in a structured manner using a format where the reasoning is transparent, traceable and reproducible. GM-VV supports this by means of a V&V argumentation structure approach. This approach can be implemented in various manners. One implementation is a V&V goal-claim network developed by the REVVA projects [10],[11]. Goal/ Claim /Argument structures originate from the Safety community and where, aside that domain, investigated and used by the ITOP (International Test Operations) to document (simulation related) evidence. Such a network provides an information and argumentation structure rooted in both goal-oriented requirements engineering and claim-argument-evidence safety engineering principles (Figure 8).

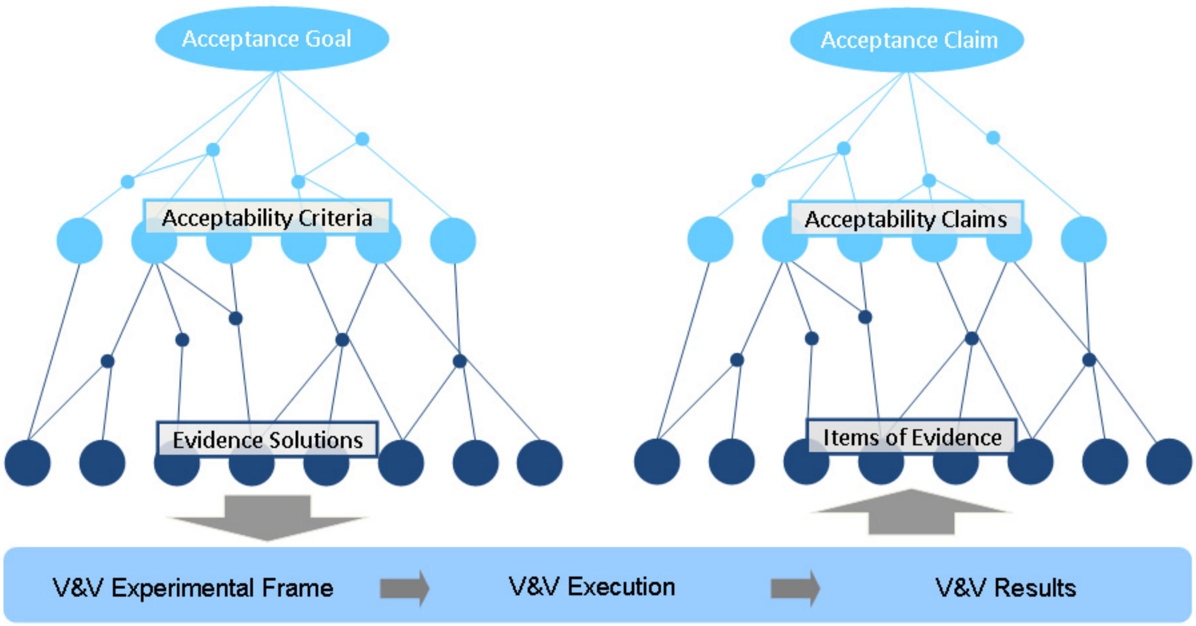

Figure 8: V&V Goal-Claim Network Implementation.

The left side of a V&V goal-claim network is used to derive the acceptability criteria from the acceptance goal; and design solutions for collecting evidence to demonstrate that the M&S asset satisfies these criteria.

Evidence solutions include the specification of tests/experiments, referent information, and methods for comparing and evaluating the test/experimental results against the referent. Collectively, they specify the design of the V&V experimental frame used to assess the M&S assets. The V&V experimental frame produces the actual V&V results that serve as the items of evidence for the right side of the goal-claim network. These items of evidence support the arguments that underpin the claims on whether or not a related acceptability criterion has been met. These acceptability claims provide the arguments for assessing to what extent the M&S asset are acceptable. This assessment eventually results in an acceptance claim for the M&S asset. The V&V goal-claim network encapsulates, manages and consolidates all underlying evidence and arguments necessary for developing an appropriate and justifiable acceptance recommendation. At the end of the V&V effort the resulting V&V Goal-Claim Network can be used to make an assessment on whether the overall V&V effort is of sufficient quality given the real world risk. This assessment must accompany the Acceptance Recommendation. Since for example insufficient resources (e.g. budget, time, skilled people, access to facilities, real world referent data) for the V&V effort may lead to weak Acceptance Recommendations for the V&V User/Sponsor risk tolerances.

The V&V argumentation structure approach can also be translated into the more traditional Systems Engineering terminology of requirements, tests and test results. A traceability matrix establishes the relationship between defined requirements (both high level and derived), acceptability criteria, test procedures, and test results. The traceability matrix provides insight into the completeness of the testing process relative to the defined requirements and identifies potential testing gaps. Figure 9 provides an example of a Traceability Matrix. This example establishes the mapping of each high level requirement (the second column) down to the test results (second to right-most column) for the implemented V&V tests. Working through the columns from left to right, the matrix documents:

A) A requirement number.

B) A high level requirement (e.g., the simulation shall be able to represent airbases and landing zones).

C) A mid-level requirement (if needed).

D) The acceptability criteria associated with the requirement.

E) A reference citation to the source document detailing how the derived requirements were defined.

F) A derived requirement (e.g., location of base, shelter options for base, logistics, damage).

G) A reference citation to the source document for the implemented tests and associated results (e.g., test plan, test report).

H) A description of the V&V test that will address the associated requirement.

I) A description of the V&V test procedures implemented.

J) Documentation of the results of the implemented V&V test.

H) Any notes related to the V&V test implantation and/or results.

A	B	C	D	E	F	G	H	I	J	K
#	High Level Requirement	Mid-level Requirement	Acceptability Criteria	Source Doc for Derived Requirement	Derived Requirement	Source Doc for V&V Test/Result	V&V Test Description	Test Procedure	Test Results	Notes
1	High Level Requirement 1.0	Mid-Level Requirement 1.1	Acceptability Criteria for Mid-Level Requirement 1.1	Source of Derived Requirements	Derived Requirement 1.1.1	Test Document(s) containing test and test result data for Derived Requirement 1.1.1	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
					Derived Requirement 1.1.2	Test Document(s) containing test and test result data for Derived Requirement 1.1.2	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
					Derived Requirement 1.1.3	Test Document(s) containing test and test result data for Derived Requirement 1.1.3	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
					Derived Requirement 1.1.4	Test Document(s) containing test and test result data for Derived Requirement 1.1.4	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
		Mid-Level Requirement 1.2	Acceptability Criteria for Mid-Level Requirement 1.2	Source of Derived Requirements	Derived Requirement 1.2.1	Test Document(s) containing test and test result data for Derived Requirement 1.2.1	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
					Derived Requirement 1.2.2	Test Document(s) containing test and test result data for Derived Requirement 1.2.2	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
					Derived Requirement 1.2.3	Test Document(s) containing test and test result data for Derived Requirement 1.2.3	Description of Test	Detailed Test Procedures	Result of test	Relevant Information
		Mid-Level Requirement 1.3	Acceptability Criteria for Mid-Level Requirement 1.3	Source of Derived Requirements	Derived Requirement 1.3	Test Document(s) containing test and test result data for Derived Requirement 1.3	Description of Test	Detailed Test Procedures	Result of test	Relevant Information

Figure 9: V&V Traceability Matrix Implementation.

As can be seen in Figure 9, multiple derived requirements can be mapped to a high level requirement. While not specifically shown, multiple tests can be associated with a single derived requirement. It is important to note that test results should not be accepted and recorded in the matrix without documented evidence (e.g., a test report) to substantiate the finding. As each test procedure is implemented, the result is documented in a formal test report. The completed tests are assessed to determine if all of the test procedures needed to address a specific requirement have been implemented and whether the test result shows that the required capability has been shown to exist and/or the level of simulation representation meets the referent with the tolerance bounds of the associated acceptability criteria. Test results for derived requirements are rolled up to determine whether or not a high-level requirement is met, partially met or not met.

5.6 V&V Management and Organizational Aspects

To facilitate V&V efficiency and quality, the V&V effort, as presented in the previous two paragraphs, should be executed in an organized way inside the V&V world. Therefore, GM-VV identifies three organizational levels where V&V efforts can be considered.

The first level is the V&V technical level, which concerns all technical aspects of a V&V effort necessary to develop and deliver an acceptance recommendation for an M&S asset. Among others, such technical aspects comprise establishing a referent, the design of experiments (DOE), and the selection of application or domain specific V&V tools and techniques [9],[4].

The second level is the V&V project level, which concerns all managerial aspects related to the proper application of all the technical aspects of a V&V effort. The *V&V Project* is a managed project, which addresses V&V planning in terms of cost, schedule and milestones. It aids in checking the V&V progress against this planning, and selecting corrective actions when needed. This V&V project could be a separate project alongside the M&S project of which the M&S asset is part of, or be an integral part of this M&S project itself (e.g. subproject, work package). A separate V&V project is relevant in the case where a level of independence must be established between the M&S development team and the V&V team (i.e. independent V&V).

The third level is the V&V enterprise level. This level defines an organizational structure, the *V&V Enterprise*, which establishes, directs and enables the execution environment for V&V projects (i.e. permanent V&V organization). At the enterprise level GM-VV distinguishes between *V&V Client* and *Supplier* entities (e.g. organizational unit and company). A V&V Client entity is an organization that acquires V&V services and products. The aforementioned *V&V User/Sponsor* is a role inside this entity. A V&V Supplier entity is an organization that provides V&V services and products. Within the V&V supplier various roles can be identified such as the *V&V Leader* and *V&V Implementer*. A single person inside such an organization can fulfil one or more of these roles. The V&V effort carried out by the V&V Supplier is based on a *V&V Agreement* between the V&V Client and Supplier.

GM-VV applies the memory concept (Figure 10) to the V&V project and enterprise levels. A memory combines an information and knowledge repository and a community of practice (i.e. human resources). The *V&V Project Memory* provides the means to manage information and knowledge produced and used during the life-time of an individual V&V project. The *V&V Enterprise Memory* retains information from past and current V&V projects to support the cost-effective execution of future V&V projects. Examples of such information include M&S technology or domain specific recommended practices, acceptability criteria, V&V goal-claim network design patterns, V&V tools and techniques.

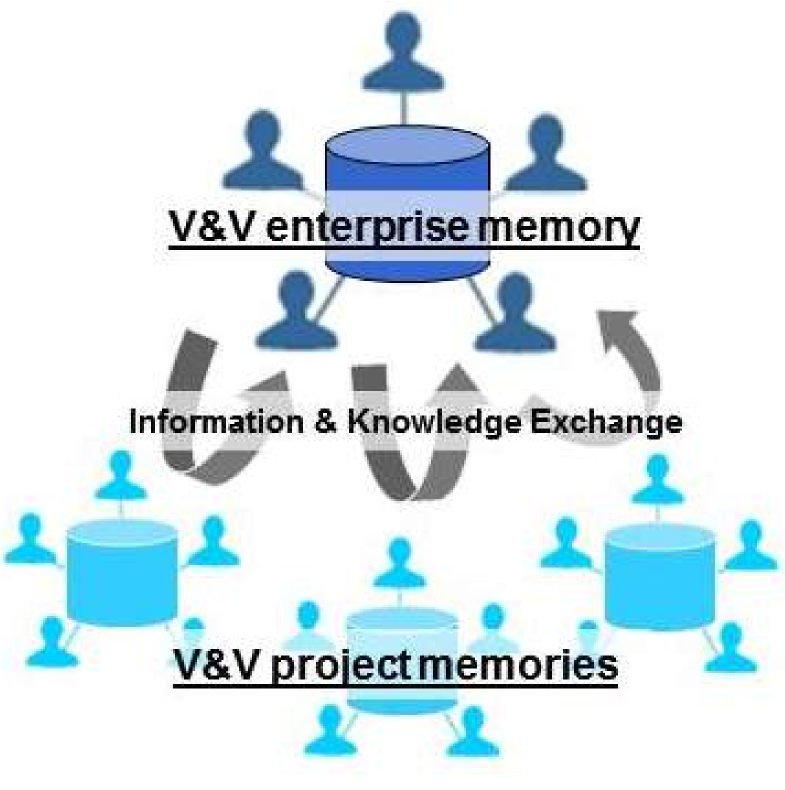

Figure 10: V&V Project and Enterprise Memory Concept.

5.7 V&V Levels of Independence: Acceptance, Certification and Accreditation

An independent V&V (IV&V) authority is often described as an organization or a person that is employed to conduct V&V, independent of the developer's team or organization. The need for IV&V is mostly driven by:

- Risks and liabilities taken by the V&V User/Sponsor's acceptance decision.
- Level of trust the V&V User/Sponsor has in the M&S developer.
- Authoritative policies and regulations that may demand independent V&V for the M&S intended use.
- Lack of specialist skills, tools and techniques by user, sponsor or developer to perform V&V.

In this context, the terms "certification" and "accreditation" are often used. Certification is the process of providing a written guarantee that a (M&S) system is acceptable for operational use [IEEE Std 100-2000]. Accreditation has two connotations. Accreditation is the official certification that a (M&S) system is acceptable for use for a specific purpose, as used by the US DoD. This meaning of the term of accreditation is the one that is part of the commonly used acronym, VV&A, which stands for verification, validation and accreditation. This acronym has a specific meaning within the US DoD M&S and decision maker community, since it integrates V&V effort within their formal acceptance decision process.

In practice however, it is highly incumbent upon the V&V User/Sponsor acceptance decision needs and complexity of the M&S system as to which parts and to what extent V&V should be conducted in an independent manner. Therefore, the GM-VV adopts a sliding scale of independence for V&V (Figure 11), which can be selected accordingly to match the V&V needs. The justification and selection of a proper level of independence is supported within GM-VV through the use of the V&V argumentation structure. Within this sliding scale for independent V&V, certification and accreditation can be located in the right part of the scale.

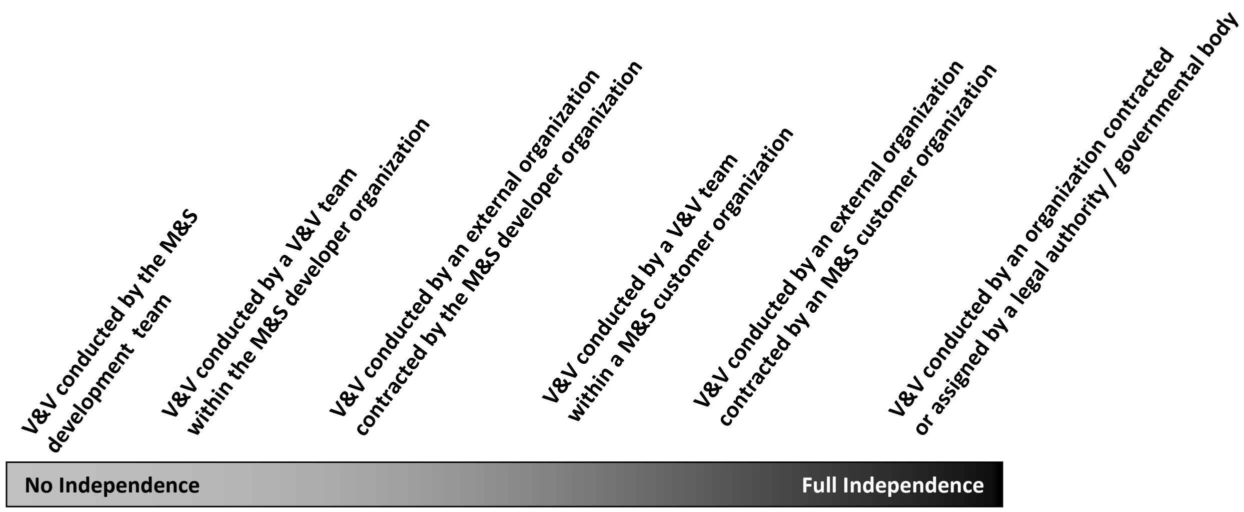

Figure 11: Levels of Independent V&V.

6.0 CONCLUSION

This paper presented the rationale for the GM-VV and its development, an overview of the technical framework, and the basics concepts part of this framework. The GM-VV concepts presented in this paper are based upon and discussed in much more detail in the original GM-VV Vol 1. and Vol 2, and the original papers of the author and other members of the GM-VV product development group. The interested reader is referred to these documents as listed in the reference section.

7.0 REFERENCES

[1] SISO, *"GM-VV Vol. 1:Introduction and Overview"*, SISO-GUIDE-001.1-2012, October 2012.

[2] SISO, *"GM-VV Vol. 2: Implementation Guide"*, SISO-GUIDE-001.2-2013, June 2013.

[3] SISO, *"GM-VV Vol. 3: Reference Manual"*, SISO-REF-039.3-2013, October 2013.

[4] US DoD, DMSO, "Recommended Practice for VV&A", Build 3.0, 2006.

[5] UK MoD, "A Generic Process for the Verification & Validation of Modeling and Simulation & Synthetic Environments Systems", DEF STAN 03-44 Issue 2, 31 March 2008.

[6] Australian DoD, DSO, "Simulation Verification, Validation and Accreditation Guide", Canberra, 2005.

[7] IEEE 1516.4-2007, "Recommended Practices for Verification, Validation and Accreditation of a Federation – An Overlay to the HLA FEDEP", December 2007.

[8] IEEE 15288, "Systems and Software Engineering – System Life-Cycle Processes", January 2008.

[9] Roza Z.C., "Simulation Fidelity Theory and Practice: a unified approach to defining, specifying and measuring realism of simulations", Delft University Press Science. December, 2004.

[10] Roza M, J. Jacquart, C. Giannoulis. Common Reference Framework for Validation, Verification and Accreditation for Simulation, REVVA-2 Reference Manual. Report: Europa 111-104. March 2009.

[11] Jacquart J., J. Voogd J., D. Brade, C. Yi, WEAG THALES JP11.20 (REVVA-1) Results and Perspectives. 2005 European Simulation Interoperability Workshop, June 2005, Paper 021.

[12] NMSG 073, Final Report on the Generic Methodology for Verification and Validation (GM-VV) to Support Acceptance of Models, Simulations and Data, NATO STO Technical Report TR-MSG-073, November 2013.

[13] Roza M, GM-VV Part 2: Processes, Information Artifacts and Roles, NATO STO LS123, April 2013.

[14] Voogd J, GM-VV Part 3: Implementation Guidance for M&S Projects and Organizations, NATO STO LS123, April 2013.