Verification, Validation and Accreditation (VV&A) – Leveraging International Initiatives

Major A.J. Masys
DND Synthetic Environment Coordination Office
Canadian Forces Experimentation Centre
National Defence Headquarters
Major-General George R. Pearkes Building
Ottawa, Ontario, K1A 0K2
Canada

masys.aj@forces.gc.ca

Dr. M. L. van Emmerik
Training & Instruction department
TNO Defense, Security and Safety
PO. Box 23, 3769 ZG
Soesterberg
The Netherlands

martijn.vanemmerik@tno.nl

LTC (Ret) Pierre BOUC
ALEKTO Etudes & Conseil
37, Traverse de la Serviane
13012 Marseille
France

alekto@wanadoo.fr

ABSTRACT

With the advent of complex coupled systems and the evolutionary introduction of new technology, the application of Modelling and Simulation (M&S) activities has increased throughout industry, academia and military domains. M&S has flourished as an enabling technology facilitating effective training opportunities from procedural training to full mission rehearsal and has been instrumental in helping decision makers take better account of the complexity, dynamics and uncertainties that pervade modern warfare. The application of M&S within the training domain brings with it an inherent risk associated with the danger of using erroneous or unsuitable models and simulation results.

Verification and Validation (V&V) of models and simulations are intended to ensure that only correct and suitable results are used thereby facilitating risk management within the training domain. To address the complexities associated with VV&A, a coordinated effort among various international VV&A working groups such as: NATO Modelling and Simulation Group (NMSG) 19; Simulation Interoperability Standards Organization (SISO): Verification, Validation & Accreditation Overlay to Federation Development Product Development Group (PDG); REVVA 2; and SISO Generic Methodology for VV&A in the M&S domain has been formulated. This paper introduces the coordinated effort of these working groups and how they contribute to the understanding, formalization, and evaluation of the quality of training.

Masys, A.J.; van Emmerik, M.L.; Bouc, P. (2006) Verification, Validation and Accreditation (VV&A) – Leveraging International Initiatives. In *Transforming Training and Experimentation through Modelling and Simulation* (pp. 10-1 – 10-10). Meeting Proceedings RTO-MP-MSG-045, Paper 10. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

1.0 INTRODUCTION

Increasingly, M&S is being exploited as an enabling technology to support tactical, operational and strategic objectives within the military domain. The use of M&S provides those within the military domain with a powerful and resource-efficient medium for:

- Exploration of doctrinal alternatives;
- Capability management;
- Alternative and complementary operational training;
- Investigation of leading-edge technologies; and
- Effective support of the acquisition process.

In support of training applications, M&S has proliferated across the services both nationally and internationally thereby facilitating joint training opportunities. Training simulation is characterized by the presence of a human in the loop and can be used for applications differing in *nature* from procedural / cognitive tasks (such as command and control tasks) to perceptual motor tasks (such as car driving), and *number* from single user (e.g., aircraft) to distributed team (network centric). Some examples include learning to drive a (passenger) car in a driving simulator, training command and control procedures in a simulated control room of a frigate, and large scale training exercises with ground troops, aircraft, marine units, and coordinating staff. The current trend in training simulation is towards these large teams, operating in a distributed fashion as in NEC or NCW.

The value of training simulation has been proven in different settings. As a result of the technological developments one can, for example, experience a realistic flight in a simulated fighter aircraft. For aircraft training at all levels simulation has become indispensable to the extent that it is now inconceivable that someone becomes a civil or military pilot without ever having experienced a simulated flight during some phase of training. In fact, when it concerns conversion training, it has been possible for over fifteen years to get licensed without training in the real aircraft - this is what Powell [1] calls 'zero aircraft flight time'.

Benefits of using simulators for training are often expressed in terms of cost (reduction), availability of the training system (as opposed to the operational equipment), and increased safety [e.g., 2,3,4]. Apart from that, there are also more didactical reasons to put forward. In a computer generated environment, the instructional designer and instructor have the flexibility to tailor reality to optimally facilitate the learning process. However, despite their widespread use, simulators are not often used to their full capacities. As the following examples illustrate, there has been a continuing lack of attention for their didactical aspects.

In the evaluation of a complex simulator system [5] the customer wanted to know if their simulator could be used for training (besides its proven testing capabilities). A quick scan of the available reference material allowed the conclusion that the main bottleneck for training was the lack of a training program (!) and proper instructor facilities. Until that time however, the focus of the customer had been on the visual system and the question if it would be sufficiently detailed to support performance of the trainees -in fact this was even better than necessary for the envisaged training use-.

Verstegen, Barnard and van Rooij [6] collected information on a total number of 39 simulator facilities throughout Europe. In their analysis of the current use of training simulators it was concluded that the possibilities of simulators were not used to their full extent. In particular facilities for instructor support, provision of feedback, and registration and assessment of performance were found to be either poor or lacking in most of them.

In yet another study the effect of simulator training for Leopard II tank drivers was found to be detrimental to performance on the operational system that is, negative transfer occurred [7]. Although this was attributed mainly to the physical aspects of the simulator (visual-, and motion systems) it was also concluded that the instructor facilities were poorly designed. Korteling [8] showed that the instructor console lacked some of the necessary facilities for provision of feedback and measurement of performance. Furthermore, it had a poor ergonomic design, and computer generated output with regard to trainee performance was not easily interpretable. Therefore, instructors did not know what information they had to use when providing feedback to their trainees. After a drastic revision of the simulator including changes to the motion- and visual systems, changes in the training program, and redesign of the instructor console and performance measurement and feedback system the simulator yielded positive transfer. These examples show that an efficient simulator based training system does not rely solely on fidelity. In fact, it is essential that content and form of training and instruction match with the needs of the trainees at each stage of their development. This can be achieved by cooperation with educational scientists and instructional designers during design and development of the simulation.

As the impact of the application of M&S within the training domain transcends the 'virtual world', the danger of adverse effects associated with erroneous or unsuitable models or simulations results also increases. In order to address this issue, the application of Verification and Validation (V&V) to the models, simulations and results is required to ensure that only correct and suitable models and simulations results are used. As discussed in detail by Brade [9], "The main driver for the V&V of models or simulation results is the risk incident to their application…. Simulation results must only be used, if they are sufficiently credible with respect to the impact of their use, and the influence of the simulation results in comparison to other non-M&S influences ("conventional" information) " [9]. The international community recognizes the requirement for VV&A and has undertaken a number of coordinated and collaborative activities to provide guidance and 'standardization' with regards to the application of VV&A for M&S. These initiatives include:

1. *NATO Modelling and Simulation Group (NMSG) 19 Task Group (TG) 016 on "Verification, Validation and Accreditation (VV&A) of Federations. This initiative is coordinated with the Simulation Interoperability Standards Organization VV&A Product Development Group (SISO VV&A PDG).*

2. *The REVVA 2 project objective is to develop and produce a set of documents, which will be proposed as a European common VV&A methodological framework for Verification, Validation and Accreditation of data, models and simulations.*

3. *The primary objective of the SISO Generic Methodology for Verification and Validation (GMV&V) Study Group (SG) is to provide a path for the creation of an internationally accepted VV&A standard complementing the current SISO VV&A PDG and other international VV&A initiatives.*

4. *The ITOP (International Test Operations Procedure) is a 4 nations (US-FR-UK-GE) initiative that focuses on standards for test operations. Part 7 focuses on simulation and V&V.*

These initiatives provide the impetus for this paper reflecting the collaborative VV&A work being conducted internationally.

2.0 VERIFICATION, VALIDATION AND ACCREDITATION

The application of M&S within the training domain brings with it an inherent risk associated with the danger of using erroneous or unsuitable models and simulation results. Viewing the simulation user as a decision maker, Harmon [10] argues that a "[*s]imulation output provides information that shapes the trainee's decision making with the goal of honing their skills, thereby improving both the decision correctness and response time. If the simulation provides incorrect information then the trainee's skill set may divert from the desired state and they may be inadequately prepared to deal with similar situations in* *the physical world. For example, a pilot trained with a simulation that does not adequately represent the behaviour of an aircraft under stall situations may not respond correctly or quickly enough to recover from an actual stall when it occurs*" [10]. Hence the requirement for a process to evaluate the credibility (confidence and correctness) [11]. Credibility (see figure 1 [9]) is a key requirement with respect to the use of M&S, particularly when the M&S application is associated with the training domain, shaping decision making and problem solving. Important characteristics of credibility are the accuracy of the simulation with respect to its intended use; the correctness associated with the level of confidence that the M&S data, algorithms, are sound, robust and properly implemented; and that the capability of the simulation matches that which is required for the specified application. VV&A activities are thereby performed to establish this credibility.

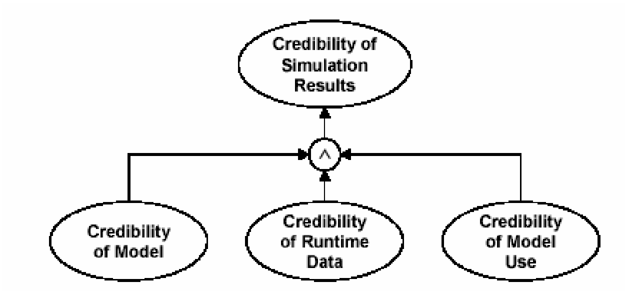

Figure 1: Credibility of simulation results

Associated with the credibility of models and simulations is the revised adage from 'garbage in, garbage out' to 'Garbage in, Hollywood out', whereby decision makers may be 'unduly influenced by state-of-the-art animation and 3D graphics that make the simulation appear more realistic than the underlying data and algorithms suggest" [12].

As articulated by Sargent [13], "Simulations models are increasingly being used in problem solving and in decision making. The developers and users of these models, the decision makers using information derived from the results of these models, and people affected by decisions based on such models are all rightly concerned with whether a model and its results are correct. This concern is addressed through model verification and validation". Tullos-Bank [14] views VV&A as "three interrelated but distinct processes that gather and evaluate evidence to determine the simulation's capabilities, limitations, and performance relative to the real-world object that it simulates, based on the simulation's intended use. The goal of VV&A is to assist the user in making an informed and independent judgment in regards to the credibility of Models and Simulations (M&S) being used in a specific program or project of interest to the user". Various definitions of VV&A exist throughout the M&S, Software and Systems Engineering field reflecting particular nuances between the applications. From the M&S perspective, Verification is defined as the process of determining that a model implementation and its associated data accurately represent the developer's conceptual description and specifications. Verification evaluates the extent to which the model or simulation has been developed using sound and established software and system engineering techniques [15]. Validation is defined as the process of determining the degree to which a distributed simulation is an accurate representation of the real world from the perspective of the intended use(s) as defined by the requirements. Validation also refers to the process of determining the confidence that should be placed on this assessment [15]. Accreditation is the official certification that a model, simulation, or federation of models and simulations and its associated data are acceptable for use for a specific purpose.

3.0 INTERNATIONAL VV&A INITIATIVES

3.1 NMSG 19/TG 16 Verification, Validation and Accreditation to FEDEP

Distributed simulations provide an architecture that facilitates opportunities to interconnect multiple simulations in support of joint training objectives. This can be realized through the advent of High Level Architecture (HLA) whereby any number of physically distributed simulation systems can be brought together into a unified simulation environment to address the needs of new applications [16]. HLA facilitates interoperability and reuse thereby leveraging the strength of individual simulations that may be geographically displaced and/or be legacy systems. The application of Verification and Validation (V&V) of models and simulations is intended to ensure that only relevant, correct and suitable models and simulation results are used. Although individual models may have been previously V&V-ed, their integration within an HLA framework may result in a simulation behaviour that yields inaccurate results thereby creating a hazardous condition where the use of these invalid results could result in negative training. Recognizing the inherent risks associated with the development of distributed simulations, a working group under the auspices of the NATO Modelling and Simulation Group (NMSG) 19 and SISO PDG was formed to develop a VV&A overlay for the HLA Federation Development and Execution Process (FEDEP) in accordance with IEEE 1516.3 [17].

HLA was developed by the Defence Modeling and Simulation Office (DMSO) of the Department of Defence (DoD) to meet the needs of defence-related projects that require a distributed simulation architecture. HLA is a software architecture for creating computer simulations out of component simulations. It provides a general framework within which simulation developers can structure and describe their simulation applications. Based on the premise that no simulations can satisfy all uses and users, HLA facilitates reuse and interoperability such that an individual simulation or set of simulations developed for one purpose can be applied to another application. Fundamental to this application is the conceptualization of the federation: a composable set of interacting simulations [19]. The development of federations within an HLA is enabled through the FEDEP, a systems engineering process. The purpose of the FEDEP is to "describe a generalized process for building and executing HLA federations. It is not intended to replace the existing management and systems design/development methodologies of HLA user organizations, but rather to provide a high-level framework for HLA federation construction and execution into which other systems engineering practices native to each individual application area can be easily integrated. In addition, the HLA FEDEP is not intended to be prescriptive, in that it does not specify a "one size fits all" process for all HLA users. Rather, the FEDEP defines a generic, common sense systems engineering methodology for HLA federations that can and should be tailored to meet the needs of user applications" [16]. The FEDEP construct facilitates the mapping of domain specific engineering practices, such as VV&A, to a distributed simulation architecture. "This mapping provides a viable basis for more detailed "how to" guides for constructing HLA federations specifically within that domain. Such mappings are commonly referred to as FEDEP "overlays" [14]. Within the federation environment, users must have assurance that models and simulations will perform as expected and can adequately support the intended purpose of the activity. A disciplined approach to VV&A of a federation can assist in lowering development and integration risk while greatly enhancing the credibility of the simulations. Such considerations drive the need for a formal, systematic, disciplined approach to the verification, validation, and accreditation of federation simulation activities. The VV&A overlay has been designed to apply across a wide range of functional applications. The purpose of the VV&A overlay is to provide a more detailed view of the VV&A processes implied by the FEDEP. Currently, these processes represent the best practices available to the VV&A community. The VV&A overlay is a tailorable process and is offered as guidance to all participants in FEDEP activities. The VV&A overlay identifies and describes the recommended VV&A processes that should be followed to assure the acceptability and utility of federations for particular intended uses. In the Reference Standard IEEE 1516.3, the FEDEP consists of seven steps and each step itself consists of activities. Every item is depicted as a list of bullets. Often a bullet contains a list of sub bullets. In the VV&A overlay a similar approach is used, in that each sub phase/activity of the FEDEP is considered from the VV&A point of view. Mapping to the 7 steps of the FEDEP, the VV&A Overlay details specific activities and product requirements to facilitate a VV&A evaluation of the simulation, figure 2.

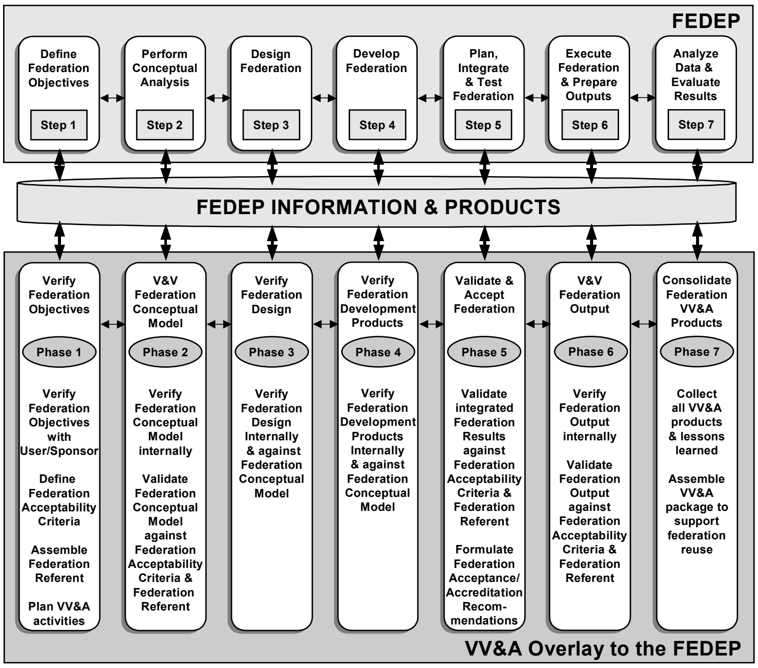

Figure 2: VV&A Overlay

3.2 REVVA 1/ REVVA 2

The REVVA (Referential for VV&A) is an European initiative which has been split into two projects:

- REVVA 1 (2004 – 2005): the THALES JP 11.20 has been conducted under the WEAG ; FR (lead), NL, DK, SW and IT were involved.
- REVVA 2 (2006 – 2009): the EUROPA 111.104 program is conducted by an international consortium composed of FR, NL, DK, SW, UK and CN.

If the only goal of REVVA 1 was to establish the basis of a European VV&A framework, REVVA 2 has two objectives: to produce a comprehensive VV&A methodology shared by the nations involved; to lead the SISO study group (GM-V&V) intended to produce a V&V standard thereby facilitating harmonization with other international initiatives.

REVVA 2 is based on concepts and principles stemming from the REVVA 1 outcomes:

- The validity of a simulation is not an inherent attribute of it; it is related to the intended purpose,
- The acceptance of a simulation is a project stakeholder decision, related to the operational objective,
- The VV&A methodology is based on a generic process which has to be tailored and instantiated to produce the effective method fitted to the specific case which is to be treated,
- The REVVA methodology is product oriented; the objective is not to apply some procedure, but to produce the relevant V&V products which contribute to a good decision making.

Some key points are to be taken into account:

- The VV&A activity has to start as early as possible,
- An important point is the definition of the different actors of the simulation project and the determination of their needs,
- Four worlds are concerned by a simulation project: the real world (where the operational objective is), the problem world (where the expected results are defined), the model world (a set of conceptual solutions), the enabled simulation product. These worlds are not to be confused while performing VV&A activities; each one has specific questions … and the VV&A process has to provide an adequate and relevant answer to each one.
- The technique of the TOA (Target Of Acceptance) allows, from the very beginning, to examine the question and the problems from the different worlds and to produce a set of acceptance criteria that will lead to the effective V&V implementation and realisation.

The REVVA 2 project (realigned and harmonized with the SISO process and products) objective is to develop and produce a set of documents, which will be proposed as a common VV&A methodological framework for VV&A of data, models and simulations. REVVA 2 will deliver four major contributions that will be used as draft documents for standardization. These documents include the following:

- The **User's manual**; this will guide users through the VV&A effort and clarify their responsibilities by explaining how to apply the methodology in practice. It describes, for example, the activities to be performed, the products to be produced, the interactions that take place among those involved, the flow of products, and how to tailor the methodology to the specific needs of the M&S project.
- The **Reference manual** documents the underlying concepts of the methodology, including the foundations of the chosen terminology, the explanation of the dependencies between activities and products, their meaning for the VV&A endeavour, and the rationale for their creation, tailoring (why and what) and execution. The reference manual is referred to whenever a deeper understanding of the methodology is required.
- The **VV&A Recommended Practices Guide** is a document devoted to an audience of M&S and V&V project leaders but without entering in the description of technical details or technical solutions.
- The **Technical Notes** document is the most detailed technical document of the VV&A methodology. It is used by V&V executioners for a very detailed knowledge of tools and techniques recommended by the methodology and for good practices of their uses.

3.3 SISO Generic Methodology for VV&A in M&S

The primary objective of the GM V&V SG is to provide a path for the creation of an internationally accepted VV&A standard complementing the current VV&A PDG efforts (dedicated to an Overlay of the FEDEP) and in consistency with the VV&A PDG efforts and other existing developments. The final objective is to provide the international community with a methodology that not only embraces a wide variety of M&S products but also may provide a future common basis for the simulation community through the GM V&V product.

As shown in figure 3, the GMV&V is harmonized within a greater VV&A initiative including Validity Process Maturity Model (VPMM), REVVA, VV&A PDG and ITOP.

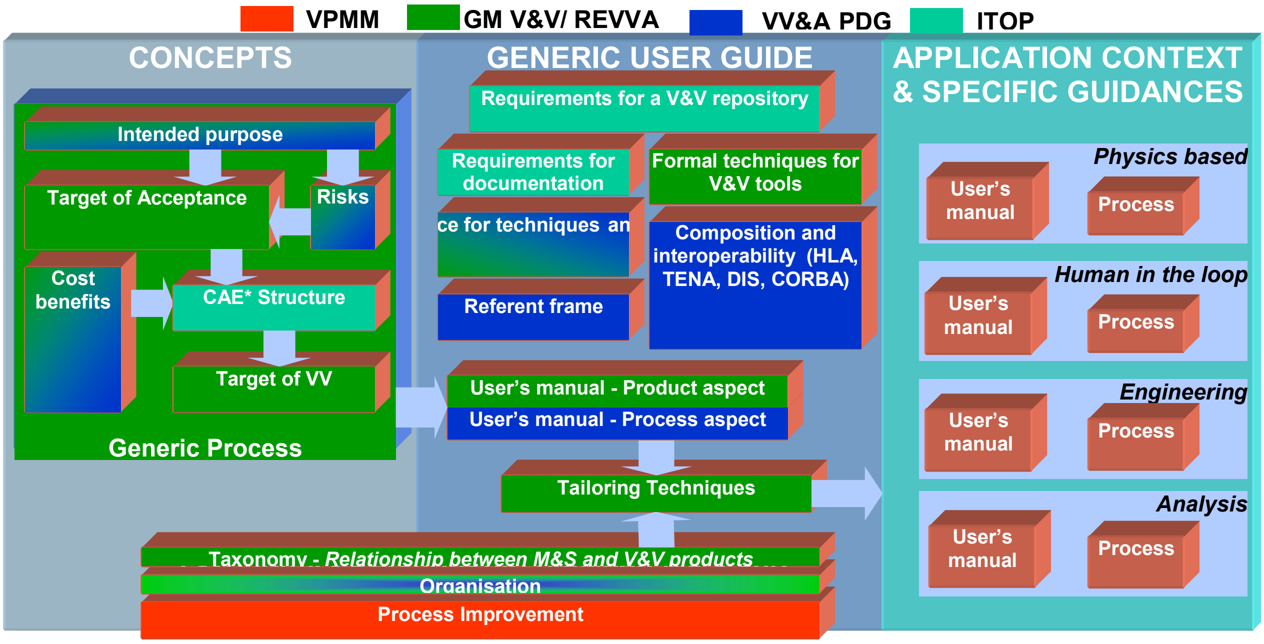

Figure 3: Unified View of International VV&A initiatives

4.0 CONCLUSION

"Given that M&S users depend on the information provided by their M&S assets as part of their decision making process, it is important they understand the limitations of that information and the risks involved in using their M&S assets. Models and simulations are abstractions of the real world and the usefulness of the information that they produce is limited by the nature of these abstractions" [15]. This is particularly relevant to the application of M&S within the military training domain whereby incorrect information garnered from M&S can negatively affect the decision making process and thereby can lead to inappropriate or incorrect actions that adversely affect their real world circumstance. The harmonization of international VV&A initiatives into a coherent and coordinated body of knowledge will seek to advance the effectiveness with regards to the application of M&S within the training domain and thereby support NATO's transformation process and national capabilities.

5.0 REFERENCES

[1] Powell, A. (1990). What price the instructor? In 1st ITEC proceedings (pp. 256-259). Birmingham, UK.

[2] Hays, R. T., & Singer, M. J. (1989). Simulation Fidelity in Training System Design. New York: Springer-Verlag.

[3] Patrick, J. (1992). Training: research and practice. London: Academic Press Limited.

[4] Stanton, N. (1996). Simulators: a review of research and practice. In N. Stanton (Ed.), Human Factors in Nuclear Safety . London: Taylor & Francis.

[5] Emmerik, M.L. van. (2003). Evaluation of the training facilities for ERA operators [In Dutch: Evaluatie van de trainingsfaciliteiten voor ERA operators]. (report TM - 03 - C011). Soesterberg, TNO-Human Factors.

[6] Verstegen, D. M. L., Barnard, Y. F., & Rooij, J. C. G. M. v. (1999). The Use of Simulators for Training and Instruction. (Report No. TM-99-B003). Soesterberg: TNO-HFRI.

[7] Breda, L. van, & Boer, J. P. A. (1988). Validation Study Driving Simulator Leopard II [In Dutch: Validatiestudie Rijsimulator Leopard II]. (Report No. Memo IZF 1988-M7). Soesterberg: TNO-Institute for Perception.

[8] Korteling, J. E. (1991). Ergonomic Design of the Instruction Consoles for the Leopard 2 and YPR-765 Driving Simulators. (Report No. IZF 1991 A-42). Soesterberg: TNO-IZF.

[9] Brade, D. (2003). A generalized process for the verification and validation of models and simulation results. Dissertation at the Fakultät für Informatik Universität der Bundeswehr München, October 2003.

[10] Harmon, S. (2005). "A Conceptual Model of Simulation Use Risk and Its Linkage to VV&A", Report on Initial Concepts for DMSO VV&A Risk and Uncertainty Driven VV&A Research Project, 27 November 2005.

[11] Brade, D. and Köster A. (2001). Risk-based Validation & Verification Levels Definition. Proceedings of the European Simulation Interoperability Workshop, London, Simulation Interoperability Standardization Organization.

[12] Roman, P.A. (2005) Garbage in, Hollywood out! Proceedings SimtecT, 2-6 May, Sydney Australia.

[13] Sargent, R.G. (2001) Some Approaches and Paradigms for Verifying and Validating Simulation Models. Published in the Proceedings of the 2001 Winter Simulation Conference. Pg 106.

[14] Tullos-Banks, H.L., Parker, C.T., Collins, K.W. (2005). Verification, Validation and Accreditation of Federations. Published in the Proceedings of the Spring Simulation Interoperability Workshop (SIW): 05S-SIW-022, San Diego, Ca. Apr 2005.

[15] IEEE 1997: Recommended Practice for Distributed Interactive Simulation – Verification, Validation, and Accreditation, IEEE Standard 1278.4.

[16] IEEE 2003**:** IEEE Recommended Practice for High Level Architecture (HLA) Federation Development and Execution Process (FEDEP). Pg iii. IEEE Standard 1516.3.

[17] IEEE 2003**:** IEEE Recommended Practice for High Level Architecture (HLA) Federation Development and Execution Process (FEDEP).IEEE Standard 1516.3.

[18] DMSO (1999): High Level Architecture. FEDEP Model Version 1.5. Dec 8 1999. Retrieved from: https://www.dmso.mil/public/library/projects/hla/guidelines/fedepv15.pdf

[19] RTO technical Report 50: NATO HLA Certification. 2002. pg 1

[20] Morse, K.L., Lutz, R. (2005) The XMSF Profile Overlay to the FEDEP. Published in the Proceedings of the Spring Simulation Interoperability Workshop (SIW): 05S-SIW-005, San Diego, Ca. Apr 2005.

[21] Harmon, S.Y., Gross, D.C. & Youngblood, S.M. (1999). "Why Validation?" Proc. Of the 1999 SCS Summer Simulation Conference, Chicago, Il, July 1999.np.