ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Computer Science 16 (2013) 878 – 887
Procedia
Computer Science

Conference on Systems Engineering Research (CSER'13)
Eds.: C.J.J. Paredis, C. Bishop, D. Bodner, Georgia Institute of Technology, Atlanta, GA, March 19-22, 2013.

The conundrum of verification and validation of social science-based models

Heidi Ann Hahn*

Los Alamos National Laboratory, PO Box 1663 MS F696, Los Alamos, NM, USA

Abstract

In the systems engineering lexicon, definitions for the terms "verification" and "validation" are settled; consistent with the definitions promulgated by the Department of Defense (DoD) [1]; and quite distinct from one another. Verification confirms that all elements of the system meet technical requirements (the product was built right). Validation confirms that the realized system complies with stakeholder requirements (the right system was built). The distinction becomes blurred, however, when one considers verification and validation (V&V) of social science-based models and simulations. Unlike physics-based models, the theoretical underpinnings of Human, Social, Cultural, and Behavioral (HSCB) or other social science models are not readily verified through observation of real-world events or empirical testing. As a result, the theoretical claims on which the models are built are often contested. As noted by Lustick and Tubin [2], when experts do not agree on what the right thing is, determining that the model is built right cannot be separated from tests of whether the right thing has been built. Because systems engineers may encounter social-science based models either in the context of system design or verification, where they may be used as substitutes for human operators or users, or when they are components of a physical system, as is the case when HSCB models are embedded into enhanced persistent surveillance systems for military or intelligence applications, it is important that they understand the limitations and controversies surrounding V&V of these types of models. In this paper, the literature on V&V of models is reviewed, with an emphasis on social science models and some recently developed constructs for their verification and validation. Future directions for social science-based model development and V&V are briefly outlined.

* Corresponding author. Tel: +1-505-665-4606; fax: +1-505-665-4181.
E-mail address: Hahn@lanl.gov.
LA-UR-12-26942

1877-0509 © 2013 The Authors. Published by Elsevier B.V. Open access under CC BY-NC-ND license.
Selection and/or peer-review under responsibility of Georgia Institute of Technology
doi:10.1016/j.procs.2013.01.092

Keywords: Verification and validation; social science-based models

1. Verification and Validation – Definitions and Overview

In the systems engineering lexicon, definitions for the terms "verification" and "validation" are settled; consistent with the definitions promulgated by the Department of Defense (DoD) **[1]**; and quite distinct from one another. Verification confirms that all elements of the system perform their intended functions and meet technical (performance and operational context) and interface requirements and constraints (i.e., the product was built right). Validation confirms that the realized system satisfies stakeholder needs, providing the correct solution to the customer's problem (i.e., the right system was built).

In the literature on models and simulations, however, the distinctions between verification and validation are not so clearly defined. Macal [3] states that verification ensures that the specification is complete and that mistakes have not been made in implementing the model. So far, so good, but he then goes on to say that verification does not ensure that the model meets a specified set of requirements. But, wasn't that part of the definition of verification? Clearly, to Macal's way of thinking, the model requirements revolve around accurately representing and correctly reproducing behaviors of the real world system. By his definitions, it is validation that ensures that the model meets its requirements in terms of the results obtained.

Sargent [4] also fails to draw boundaries between verification and validation and speaks of model validation as determining whether the model correctly represents a governing theory (whether the model was built right) and whether the model works in a "reasonable way" given its purpose. Sargent, then, intertwines verification and validation.

In the world of modelling and simulation, the distinctions between verification and validation have been collapsed into the generic problem of validation.

1.1. Discussion of Validation as it is Found in the Social Science and Theory-Based Modeling Literature

Validity, in its generic form, "refers to measuring what we think we are measuring" [5] or, in the case of models, representing what we think we are representing. Indeed, there is much is the social science literature about the validity of models and methods for establishing model validity. Thomas [6] asserts that model validation efforts must be concerned with internal validity—the extent to which the relationships between variables are represented correctly in the model (verification) – as well as external validity, or the extent to which the model outputs agree with an external entity, which may be either a real world system or another validated model (validation).

In essence, all model validation efforts attempt to establish construct validity, or the extent to which the model accurately represents a theoretical construct or characteristic. Trochim [7] writes of two types of construct validity: translation validity and criterion-related validity. Translation validity focuses on whether the model operationalization is a good reflection of the underlying constructs. Translation validity comprises face validity (the extent to which the model appears to represent accurately what it is intended to represent) and content validity (the extent to which the model sufficiently covers the domain of interest). Trochim notes that content validation assumes that the content domain can be described accurately, an assumption that is not always true in the case of social science models. This is one of the characteristics of social science-based models that becomes problematic for model V&V.

Criterion-related validity considers the performance of the operationalized model. Model performance may be measured against one or more of the following four forms of criterion-related validity [7]:

- Predictive validity – the ability of the model to predict something that it should theoretically be able to predict
- Concurrent validity – the ability of the model to distinguish between two entities that it should theoretically be able to distinguish between

- Convergent validity – the degree to which the model outputs are similar to the outputs of other models that they should be similar to (a form of external validity as described by Thomas [6])
- Discriminant validity – the degree to which the model outputs are not similar to the outputs of other models they should not be similar to

2. Comparison of Verification and Validation of Physical Systems to V&V of Theory-Based Models

Semantics aside, V&V of physical systems versus theory-based models, be they models of physical or physics-based systems or human-based or social science systems, can be thought of as existing along a "continuum of objectivity" as shown in Figure 1. Much of the difference in V&V of physical systems and theory-based models lies in the objectivity of the evidence basis used.

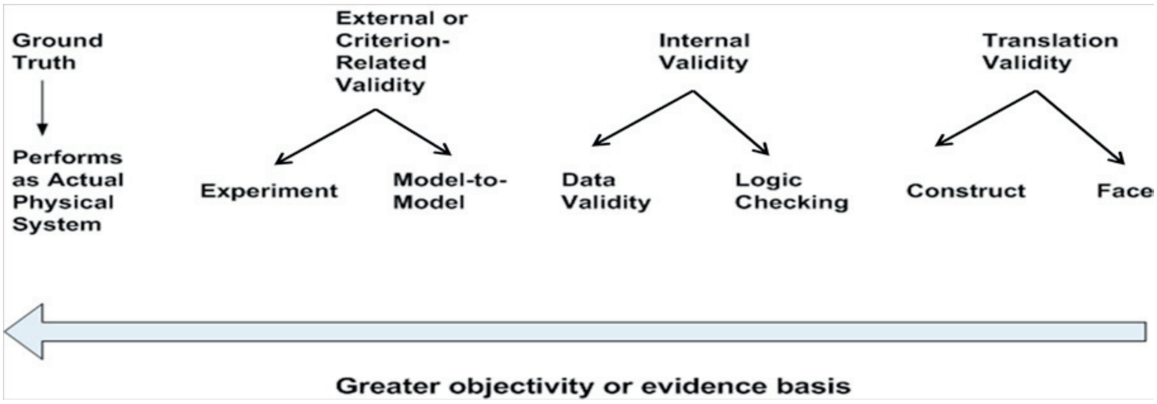

Fig. 1. V&V Continuum of Objectivity

2.1. V&V of Physical Systems

In the V&V of physical systems, there is "ground truth" against which the as-built system can be measured – it can either fly so far or it can't, it weighs less than X or it doesn't, and so on. In validating physical systems, the left-most side of the graphic shown above is the normal operating regime – we want to prove that the system meets the performance requirements. Testing – measuring the system's response to a prescribed set of conditions, which may be real or simulated, and comparing the results to a set of specified operability, supportability, or performance capability requirements – is the preferred V&V method for physical systems. When testing is infeasible, such as when tests would result in destruction of the system, analysis using experimental data, models, and simulations to show theoretical compliance of the system to specified requirements is an acceptable V&V method.

Two fundamental assumptions underpin V&V of physical systems. As stated above, first, there is the assumption that the system can be verified against "ground truth," as defined by an observable set of technical requirements (quantified statements that define how well, under what operational conditions, and/or to what degree functions must be performed) and constraints that enumerate specific parameters that the design must meet. Verification activities provide objective evidence, through testing, analysis, demonstration, and/or inspection, that the physical system performs the required functions as specified.

Second, there is the assumption that the technical requirements and specifications flow hierarchically from validated customer and stakeholder needs. In theory, then, a system that has been fully verified, proving that all requirements have been satisfied, should also be valid. In practice, it is sometimes the case that a system will have been designed and implemented perfectly, meeting all requirements, but still not provide the correct solution to the customer's problem – the verified system is not valid. Shamieh [8] provides the following example of such a situation:

> "Take, for instance, the case of a well-engineered traffic light control system. Expertly designed to control the sequence of traffic lights in a large city, this system might fail to meet its *intended* purpose – to reduce congestion by streamlining the flow of traffic – if no one bothered to study the city's typical traffic patterns and map them to system requirements for timing sequences." (p. 47)

In such cases, the problem can almost always be traced back to an incorrect, ambiguous, incorrect, or missing requirement: the system as built is not valid because the design failed to adequately account for one or more customer or stakeholder requirements.

2.2. V&V of Theory-Based Models

When considering V&V of theory-based models, however, either, or both, of the assumptions made when validating physical systems may be invalid. The option of validating against ground truth (i.e., historic data collected from a real system) is often not available to modellers. Further, as Lustick and Tubin [2] point out, theory-based models may pass verification checks, proving that the model was a correct operationalization of the design concept, but fail validation tests either because either because the design was not a valid implementation of the underlying theory or because the theory on which the model was based was not valid. "In that case, the model would have inherited the invalidity of the theory upon which it was based" (p. 6586).

Typically, V&V of theory-based models exists in the middle of the range of the objectivity continuum, taking an empirical approach, ideally by comparing the results of test cases to experimental data or, when experiments are not possible, to other validated models. Analytic techniques may also be used to show that the elements of the model are correct and are correctly integrated and to determine the model's "fit for purpose" [9].

Internal validity checks are a key element of the V&V of theory-based models. In performing internal validity checks, the modeller verifies that the right data and logic have been captured. Because inaccurate data could be a significant source of inaccuracy in the model outputs, V&V of models typically includes analyzing data for inconsistencies, incorrect formatting, and gaps (unavailability of data); estimating the effects of inaccuracies; and putting procedures in place for collecting or estimating data for which high sensitivity exists [10]. Logic checks can be done by watching the model run over a period of time and observing whether the actual model behavior conforms to the expected behavior [10].

2.2.1. Special Considerations for V&V of Social Science-Based Models

Arguably, the "gold standard" for theory-based model validation is the "Standard Model" (the theory of the physical laws governing the fundamental particles of matter and their interactions), whose predictions have matched experimental data with great accuracy. Other science-based Standard Models are also emerging – for example, advances in observing the large-scale Universe have led to a well-tested and cross-verified "Standard Model" of cosmology. In the social sciences, however, there is no Standard Model, nor is there likely to be one in the foreseeable future. Indeed, the constructs underlying most social science-based models are regularly contested. (Take, for example, the debates over theories of leadership, and whether leadership qualities are inborn or learned, that have been going on since the introduction of the "great man" theory of leadership in the late 1800's.) As noted by Lustick and Tubin [2], when experts do not agree on what the right thing is, determining that the model is built right cannot be separated from tests of whether the right thing has been built.

To complicate matters further, Ruvinsky, Wedgwood, and Welsh [11] correctly observe that the empirical approach to V&V falls short for social science-based models in the following areas:

- Experiment design – large-scale, naturalistic domain spaces pose challenges to the design and control of experiments
- Data acquisition – social phenomena tend to be described by multiple variables whose measurement may not be clearly understood

- Comparison of model results to real world data – real world data sets, either current or historical, may not be available or, if available, may not be sufficiently robust to support model validation

As a result, V&V of social science-based models relies more heavily on the right-hand end of the objectivity continuum, with internal and translation validation techniques often being the primary methods used. Construct validity, in particular, aims to address the issue of experts not agreeing on what the "right thing" is.

3. Recent Innovations in V&V of Social Science-based Models

Recent work on the validation of agent-based and Human, Social, Cultural, and Behavioral (HSCB) [a] models provide insight into how to address some of the challenges described previously. Macal and North [12] have proposed innovative methods for V&V that build upon traditional V&V techniques, placing particular emphasis on internal validation and designing novel methods for performing model-to-model comparisons. Ruvinsky, Wedgwood, and Welsh [11] have proposed a radically different methodology, which aims to perform construct validation based on analysis of the epistemological structure of the model rather than on the operationalization of the model *per se*. Each of these approaches is reviewed in some detail below.

3.1. Innovations in the Validation of Agent-based Models

Macal and North [12], addressing the issue of validation of agent-based models, have proposed innovative methods for V&V to validate EMCAS (Electric Market Complex Adaptive System), an agent-based simulation model designed to investigate the effects of electric power market restructuring and deregulation on electricity costs, reliability, and availability. As is the case with many social science-based models, the domain being modeled, a deregulated electric power market, did not exist, so comparison with real-world behavior was not an option. Thus, their task was to "establish an argument that the model produces sound insights and sound data based on a wide range of tests and criteria that 'stand in' for comparing model results to data from the real system" (p. 3). They constructed a validation framework and a rigorous process to validate the model, providing a set of resources that can be used to counter objections to the validity of the model and the results it produces. The framework comprises the following elements:

- Data validation
- Subject matter expert (SME) judgment
- Participatory simulation
- Model-to-model comparison
- Critical tests and key indicators
- Comprehensive test cases
- Invalidation exercises

Noting that data gaps or inconsistencies can invalidate model results and destroy the model's credibility, Macal and North [12] used a variety of iterative analysis techniques to validate the data inputs to the model. These included:

- Mapping and cross-referencing data definitions from different databases and converting data to common units to ensure consistency
- Assuring currency of the data
- Engaging third-parties (stakeholders) in data verification
- Including proprietary data provided by stakeholders
- Using data visualization techniques to identify data anomalies that were not readily apparent through inspection

---

[a] HSCB models are a special class of social science model intended to help the United States military forecast the behaviors of key individuals and groups in foreign operational contexts and develop effective strategies and courses of action to address them [15]. The HSCB community is particularly interested in model validity and the DoD's Office of Naval Research (ONR) has an active research program, called MESA or Model Evaluations, Selection, and Application, in this arena [13].

Data validation was an exhaustive process, taking several months and delaying progress on the main modeling effort, but paid off, as all data-related uncertainties and anomalies were resolved establishing a sound foundation for the EMCAS model [12].

Macal and North [12] used a workshop format to engage independent SMEs in evaluating the model, model assumptions, and agent behavior during design. The independent SMEs were instrumental in identifying test cases that would reveal model weaknesses during the invalidation exercises. In these exercises, multiple models runs were executed in systematic attempts to have the model exhibit unexpected behaviors, thereby invalidating the model assumptions. Unexpected behaviors became focal points for more in-depth analysis and model refinement. Macal and North [12] note that the systematic design of the invalidation exercises was key to avoiding validation bias, or the tendency to perform only those validation tests that are likely to validate the model.

Among the most novel of Macal and North's [12] validation methods was the use of "participatory simulation" (p. 4) in which real people played the roles of agents in the energy market. Results of the simulation were compared to the model results. The EMCAS results closely matched the results of the participatory simulation, validating model assumptions about agent strategies and behaviors under deregulation.

Using more traditional V&V methods, Macal and North [12] also generated test cases to evaluate model performance in replicating known system behaviors, such as replicating results for the regulated energy market; compared EMCAS against two other validated models representing special cases of EMCAS's operational parameters; and performed comprehensive testing across the full spectrum of plausible agent strategies (bounded by the assumption of rationality) and parameter settings for variables such as price, quantity, and generating capacity.

Macal and North [12] concluded that use of the validation process described above did result in the model being better accepted as a tool for answering important questions with respect to electricity deregulation, and that the process they developed is a generalizable and practical framework for agent-based model validation.

3.2. Innovations in the Validation of HSCB Models

DoD's Office of Naval Research (ONR) has recognized the challenges inherent in verification and validation of social science models and issued a Broad Area Announcement (BAA) that included tasking related to the validation of HSCB modeling software for integration into existing system architectures [13]. Responding to the ONR BAA, Ruvinsky, Wedgwood, and Welsh [11] have proposed a radically different V&V methodology, which is based on analysis of the epistemological structure of the model rather than on the operationalization of the model *per se*. They challenge the assumption that V&V methods for social science-based models should move to the left-hand side of the objectivity continuum and claim that their method "move[s] beyond viewing verification and validation as limited by empirical testing by providing mechanisms and techniques to verify and validate all aspects of knowledge/information that a model uses or produces" (p. 6595). Essentially, their goal is to develop a more robust construct validation method as an alternative to more traditional "operational" V&V.

Ruvinsky, Wedgwood, and Welsh's [11] methodology takes what they call the focal V&V point of view, in which assessment focuses on "how well a model explains the phenomenon for which it was designed" (p. 6598). The first step in their method is the decomposition of the model into its hierarchical epistemological elements, as shown in Table 1. For each tier in the hierarchy, there are computational model artifacts and data artifacts that can be examined during the course of V&V. For example, at the Data Level the data sources (e. g., specific databases) and raw data that serves as input into model parameters would be the artifacts associated with the computational model and the data model, respectively.

Evaluators then independently assess the model against a series of "metadata elements" presented in questionnaire format, assigning a numerical score or letter grade to each and noting discrepancies found. The "metatags" for V&V at each level of knowledge in the epistemological hierarchy are shown in Table 2.

Table 1. Elements of the epistemological hierarchy of a HSCB model [11, p. 6596]

Element	Definition
Social Ontology	The set of background entities and beliefs about the world that pertain to and characterize a basic structure of reality
Paradigm or Conceptual Repertoire	A grand scheme or worldview that brings to bear the basic concepts prior to claims about specific domains
Theories	Abstract statements about reality describing relationships between or among concepts
Social Model	A representation of real world system behavior based on theories and concepts
Hypotheses	Conjectures within a theory regarding the relationship of two concepts to be explored in the Social Model
Application Model	A description of how the Social Model will be refined and the concepts operationalized
Implementation Model	The equations, parameter settings, and coding rules to enable execution of the Application Model and manipulation of the raw data into model parameters
Data	Selection of specific data bases and methods to access of specific data sets from specific databases

Key:

	Conceptual Levels		Theoretical Levels		Operational Levels

Results across reviewers are analyzed for similarities and differences, and scores and evaluator comments are used to make judgments about verification and validation at the various levels of the hierarchy. Ruvinsky, Wedgwood, and Welsh [11] assert that this ability to assess, for example, the V&V of a hypothesis separate from the V&V of the specific application model of the social construct from which it is derived, not only helps in managing the V&V effort but also enables model improvements by isolating the faulty parts of the model's epistemology.

Ruvinsky, Wedgwood and Welsh [11] observe that the exercise of decomposing the model into its epistemological hierarchy can provide insight into the feasibility of performing V&V on the model, stating:

> "The difficulty of teasing out the levels of the epistemological hierarchy of a model from the documentation is indicative of the difficulty and the cost that would be incurred in attempting to perform V&V on the model. Models with easier decompositions lead to more accessible V&V assessment" (p.6598)

This suggests that, even if the utility of focal validation using epistemological hierarchies is not borne out, the method may still be valuable as a pre-screening tool when undertaking model validation efforts.

Sallach [14], who is part of the research group working on ONR's MESA program, notes that focal V&V alone is not sufficient for social science-based models because contextual factors that may appear tangential to the model often determine its robustness. His work focuses on contextual V&V, which uses epistemological hierarchies to assess the model's susceptibility to perturbations from effects not explicitly captured in the model. Although this work has not yet sufficiently progressed to provide a detailed methodology, it appears to be addressing issues that bridge a form of criterion-related validity – that the model should not be sensitive to variables not thought to be important enough to system behavior to have been included in the model – and construct validity – that the theoretical construct is accurately represented.

Table 2. Focal V&V Metatags and Questions [11, p. 6599-6600]

	Verification	Validation
Conceptual Levels	Abstract *Are components appropriately simple?* Ensemble *Are components posed at an appropriate intermediate level of specificity?* Virtualization *Are components sufficiently differentiated?*	Theoretical Provenance *How substantively persuasive are the theories used and how prevalent is scientific work based on the theories?*
Theoretical Levels	Unit of Analysis *Are the units of analysis consistent with the units at the conceptual level?* Assumptions *What proportion of assumptions are specified?* Dependent Concepts *Are outcomes of interest clearly specified?* Independent Concepts *Are causes or patterns used to explain variation explained clearly?* Intervening Concepts *How well are mechanisms linking cause and effect specified?*	Persuasiveness of Explanation *How persuasive is the strategy for testing and evaluation?*
Operational Levels	Methodologies *How appropriate are the methods chosen for evaluating claims?* Fidelity and Reliability of Operationalization *Do the measurements meaningfully reflect the construct being measured? Are the measures repeatable?* Fidelity of Abstraction *To what extent is a general claim warranted based on integration of more specific corroborated claims?* Replication *Can the results of the study be reproduced?*	Testing *How much testing has been done and was the sample selection appropriate?* Substantive Findings *Are the theoretical, empirical, and/or policy relevant findings substantively useful?*

4. Application of Social Science-Based Models in Systems Engineering

It is important for systems engineers to understand the differences between V&V of physical systems and theory-based models and to appreciate the conundrum of validating social science-based models. Systems engineers may encounter social science-based models in several contexts. First, social science-based models may be useful in the design and verification of physical systems, for testing system behavior when incorporation of a human operator or user is not possible or practical. One example of such a use is a Smart Grid project, in which the behavior of human energy consumers in a demonstration house is modeled and provided as input to the grid controllers rather than instrumenting an actual home to provide data on the occupants' energy consuming behaviors. In this context, the systems engineer must have confidence in the credibility of the model – understanding how the model was validated may help build trust in the model outputs.

In addition, social science-based models are increasingly being incorporated into physical systems as a system component intended to satisfy certain functional requirements. For example, as described in the Future Directions section below, there is interest within the DoD in embedding HSCB models into enhanced persistent surveillance systems for military and intelligence applications [13]. In such cases, the systems engineer may well find him/herself in the position of needing to validate the model as part of the larger system verification effort. Understanding the limitations of V&V of social science-based models and being aware of techniques to mitigate or overcome them should prove useful in this context.

5. Conclusions

To be accepted by decision makers, theory-based models, including social science models, must credible. Model validation is essential to credibility. Traditional, empirically-based V&V methods, however, are often not suitable for validation of social-science based models, either because there is no real world system behavior to compare the model results to or because real world data collection is intractable. Thus, researchers have been seeking alternative methods and frameworks for model validation.

In this paper, two different approaches to establishing model credibility were reviewed. In one approach [12] traditional V&V methods including testing and model-to-model comparisons were effectively augmented with interventions such as reviews by independent SMEs and participatory simulations designed to engage stakeholders in model validation and, perhaps more importantly, in exercises attempting to invalidate the model. Although developed for validation of agent-based models, these novel methods can be applied generically across a wide variety of social science-based modeling applications.

In the other [11], model verification and validation proceeds against an epistemological hierarchy of the model rather than on the model *per se*. Although in its early stages of development and still in need of additional testing, the method appears promising, especially if the research group is successful in extending the framework to contextual validation. One strength of Ruvinsky, Wedgwood, and Welsh's [11] approach is that it addresses the issue of validity or invalidity of underlying theory directly, performing V&V at both conceptual and theoretical epistemological levels.

6. Future Directions

Although there has been significant progress in the development of models for understanding, detecting, predicting, and effecting change in human behavior, there is much work still to be done. Among the highest research and engineering priorities for the HSCB and social science-based modeling community are:

- Building quantitative underpinnings for sociocultural behavior, with the goal of achieving a degree of rigor comparable to that of the scientific Standard Models
- Developing new methodologies and tools for the valid collection of sociocultural behavioral data, as well as a repository of sociocultural behavior data and ontologies
- Researching and engineering a "social radar" which would be a global and persistent indications and warning capability for detecting an monitoring relevant sociocultural behavior signatures, and integrating analytics of sociocultural behaviors with conventional and geospatial data [13]
- Integrating social science models from intersecting domains to improve the breadth of understanding of complex human systems (akin to the "Beyond the Standard Model" work taking place in physical system modeling, which is intertwining physics- and cosmology-based models)
- Developing multi-platform modeling systems (integrating game theory, systems dynamics, and agent-based modeling) that decision makers can use to explore alternative courses of action

Building a rigorous foundation for social science-based computational models, by its nature, must include establishing and building consensus on methods for V&V of such models [13]. Whether either of the approaches described above brings the community to such consensus remains to be seen, particularly because demonstrating that the models validated using these methodologies are "more credible" – which both Macal and North [12] and Ruvinsky, Wedgwood, and Welsh [11] hold out as their goal – may not be very satisfying. Rather, the challenge may be to move beyond thinking of model validation as primarily to establish "credibility" of the model and to develop metrics that begin to answer the question of whether models validated using these methods are "better" on a variety of dimensions – are the data more accurate, is the logic more internally consistent, are the models better predictors of whatever system behavior they purport to represent, etc. – to provide evidence that social science-based behavioral models are on par with the “goodness” of the science-based Standard Models.

References

[1] Department of Defense, "Key Concepts of VV&A," 2006. [Online]. Available: http://vva.msco.mil/Key/key-pr.pdf. [Accessed 23 August 2012].

[2] I. S. Lustick and M. R. Tubin, "Verification as a Form of Validation: Deepening Theory to Broaden Application of DoD Protocols to the Social Sciences," in *Proceedings of the 4th International Conference on Applied Human Factors and Ergonomics*, San Francisco, 2012.

[3] C. M. Macal, *Model Verification and Validation,* Chicago: The University of Chicago and Argonne National Laboratory, 2005.

[4] R. Sargent, "Validation and Verification of Simulation Models," in *Proceedings of the 2004 Winter Simulation Conference*, Washingston, DC, 2004.

[5] G. King, R. O. Keohane and S. Verba, Designing Social Inquiry: Scientific Inference in Qualitative Research, Princeton, NJ: Princeton University Press, 1994.

[6] C. Thomas, "Verification Revisited," in *Military Modeling for Decision Making*, Alexandria, VA, Military Operations Research Society, 1997, pp. 333-350.

[7] W. M. Trochim, "Measurement Validity Types," 20 October 2006. [Online]. Available: http://www.socialresearchmethods.net/kb/measval.php. [Accessed 2012 September 24].

[8] C. Shamieh, Systems Engineering for Dummies, Indianapolis: Wiley Publishing, Inc., 2011.

[9] G. J. Friedman, Constraint Theory: Multidimensional Mathematical Model Management, New York: Springer Science+Business Media, Inc., 2005.

[10] S. Robinson, "Simulation Model Verification and Validation: Increasing the Users' Confidence," in *Proceedings of the 1997 Winter Simulation Conference*, Atlanta, 1997.

[11] A. I. Ruvinsky, J. E. Wedgwood and J. J. Welsh, "Establishing Boaunds of Responsible Operational Use of Social Science Models Via Innovations in Verification and Validation," in *Proceedings of the 4th International Conference on Applied Human Factors and Ergonomics*, San Francisco, 2012.

[12] C. M. Macal and M. J. North, "Validation of an Agent-based Model of Deregulated Electric Power Markets," 2007. [Online]. Available: http://www2.econ.iastate.edu/tesfatsi/EmpValidACE.MacalNorth.ElectricPower.pdf. [Accessed 20 September 2012].

[13] Office of Naval Research, "BAA Announcement Number 09-026," 2009. [Online]. Available: https://www.fbo.gov/utils/view?id=ea459ff2dde3362048607957ccc830a0. [Accessed 24 September 2012].

[14] D. L. Sallach, "Contextual Validation: Concept and Application," in *Proceedings of the 4th International Conference on Applied Human Factors and Ergonomics*, San Francisco, 2012.

[15] Office of the Secretary of Defense, "Sociocultural Behavior Research and Engineering in the Department of Defense Context," September 2011. [Online]. Available: http://www.dtic.mil/biosys/files/SBRE2011.pdf. [Accessed 25 September 2012].