Verification and Validation Methodology for Computational Cognitive, Behavioral, and Social Models

Contacts: Laura A. McNamara, Exploratory Simulation Tech., *lamcnam@sandia.gov*
Timothy G. Trucano, Optimization and UQ, *tgtruca@sandia.gov*
George A. Backus, Exploratory Simulation Tech., *gabacku@sandia.gov*
Scott A. Mitchell, Computer Science and Informatics, *samitch@sandia.gov*

Sandia National Laboratories is targeting a leadership role in developing Verification and Validation (V&V) approaches for computational social science applications.

Across government and industry, computational modeling methodologies such as agent-based models, systems dynamics, and artificial neural networks are assuming a prominent role in simulations involving social, cognitive, and behavioral phenomena. As leaders make greater investments in computational simulations for policy decisions, methodologies for evaluating models will become increasingly important. Sandia National Laboratories is targeting a leadership role in developing Verification and Validation (V&V) approaches for computational social science applications. This requires understanding the unusual characteristics of V&V in computational social science, as well as its technical difficulties. We have suggested a methodological framework for addressing these challenges that builds on our previous experience in computational physics and engineering and that can be systematically incorporated into decision makers' assessments of computational social simulations.

In the United States, across government and industry, mathematical modeling and computational simulation have long been perceived as critical to decision making, in areas from nuclear weapons design and development to operations research. Until the 9/11 attacks, simulations of sociocultural phenomena were less common in policymaking circles. However, since 2001, federal agencies have invested heavily in computational modeling and simulation software to help analysts and decision makers better understand "soft" problems, like insurgency and terrorism. A 2006 article in *IEEE Spectrum* pointed out that the United States' national security enterprise is betting heavily that "computers equipped with the right software can give vital insights into the minds and motives of terrorists and the structure and critical links in their organizations" [1].

Largely lacking are comprehensive approaches to evaluating the goodness of the insights that computers provide decision makers. Sandia is particularly well placed to address this problem: Under the Advanced Simulation and Computing Program (ASC), Sandia assumed a leading role in developing verification and validation (V&V) methodologies for assessing the internal and external correctness of computational simulations. In our work, we have been leveraging the principles of the ASC V&V Program in developing a robust framework for evaluating computational simulations of social and/or psychological phenomena.

In computational science and engineering (CS&E), verification and validation refer to a suite of methodologies that assess the internal (verification) and external (validation) correctness of a model-based simulation. Most of the specific methods developed for V&V at Sandia derive from computational physics or engineering problems, where mathematical or logical proof can be combined with planned experiments to develop rigorous assessments of a code's adequacy for the decisions its developers intend it to support. Developing verification and validation approaches appropriate for the simulation techniques used in computational social science—which range from game-based training models to systems dynamics models—presents a range of challenges, from methodological to organizational. These include the development of verification methods, the difficulty of designing and conducting high-fidelity social experiments for validation, and the problem of balancing resources for simulation evaluation against equally pressing demands for model development and application in resource-constrained environments.

To address these challenges, we have focused on developing a high-level framework for simulation evaluation that incorporates a range of domain-independent concepts for verification and validation activities. We begin by defining verification and validation *as the systematic accumulation of evidence that a model/simulation is adequate for application in the intended decision context.* This means that planning for verification and validation activities begins well before coding starts, when a decision context is identified and simulation requirements specified.

In this regard, several elements of the ASC V&V Program can be leveraged to support robust evaluation. Two examples are the Predictive Capability Maturity Model (PCMM) and the Phenomenon Identification and Ranking Table (PIRT), tools used to systematically plan and implement evaluation activities. A PCMM identifies elements of the simulation and maps requirements for their technical maturity to the application of the simulation. High-consequence decisions that rely heavily on simulation results demand correspondingly robust demonstrations of maturity, which implies significant verification and validation investments. The PIRT requires that teams systematically identify the phenomena that the simulation is attempting to replicate and documents the relative importance of each phenomenon and the degree to which it is adequately understood. In doing so, the PIRT helps modeling teams prioritize V&V investments in relation to the evaluation requirements implied by the PCMM [2].

To demonstrate the cross-domain applicability of ASC-based approaches, we worked with a cognitive science modeling project to develop an evaluation-planning rubric that incorporated many ASC approaches and concepts, including the PCMM and the PIRT. This resulted in the development of an extensive evaluation approach for a particular cognitive model of human memory formation, details of which are available in our recent report [3].

References

1. Goldstein, H. (2006). Modeling Terrorists: New Simulators Could Help Intelligence Analysts Think Like the Enemy, *IEEE Spectrum* (September 2006), pp. 26-35.

2. Oberkampf, W. L., Pilch, M., and Trucano, T. G. (2007). *Predictive Capability Maturity Model for Computational Modeling and Simulation*, Sandia National Laboratories, Albuquerque, NM.

3. McNamara, L.A., Backus, G. A., Trucano, T. G., Mitchell, S. A., Slepoy, A. (2008). *R&D for Computational Cognitive and Social Models: Foundations for Model Evaluation through Verification and Validation* (Final LDRD Report), SAND 2008-6453, Sandia National Laboratories, Albuquerque, NM.

This work has been supported by SNL's Laboratory Directed Research and Development (LDRD) Program.

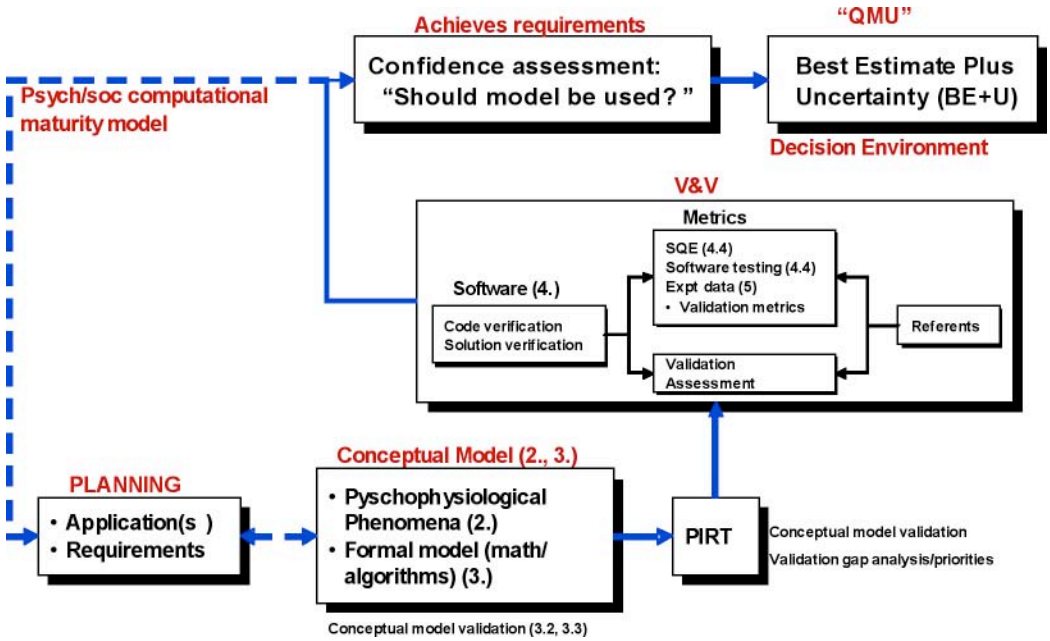

Figure 1. Adapting ASC V&V methods to psychological and social computational simulation.