Model Verification and Validation

Charles M. Macal*

Workshop on
"Threat Anticipation: Social Science Methods and Models"
The University of Chicago and Argonne National Laboratory

April 7-9, 2005
Chicago, IL

***Center for Complex Adaptive Agent Systems Simulation (*CAS2*)**
Decision & Information Sciences Division, Argonne National Laboratory

University of Chicago

The Need for V&V

- **Model verification and validation (V&V) are essential parts of the model development process if models to be accepted and used to support decision making**
- **One of the very first questions that a person who is promoting a model is likely to encounter is "has your model been validated?"**

If the answer to this critical question is *No…*

- Experience has shown that the model is unlikely to be adopted or even tried out in a real-world setting
- Often the model is "sent back to the drawing board"
- The challenge then becomes one of being able to say "yes" to this critical question

Model Verification:
## *Does the model perform as intended?*

- **Verification is done to ensure that:**
  - The model is programmed correctly
  - The algorithms have been implemented properly
  - The model does not contain errors, oversights, or bugs
- **Verification ensures that the specification is complete and that mistakes have not been made in implementing the model**
- **Verification does not ensure the model:**
  - Solves an important problem
  - Meets a specified set of model requirements
  - Correctly reflects the workings of a real world process

Practical Verification

- No computational model will ever be fully verified, guaranteeing 100% error-free implementation
- A high degree of statistical certainty is all that can be realized for any model as more cases are tested
  - Statistical certainty is increased as important cases are tested
  - In principle, a properly structured testing program increases the level of certainty for a verified model to acceptable levels
  - Exercise model for all possible cases
  - Automated testing process
- Model verification proceeds as more tests are performed, errors are identified, and corrections are made to the underlying model, often resulting in retesting requirements to ensure code integrity
- The end result of verification is technically not a verified model, but rather a model that has passed all the verification tests!

The Model Validation "Gold Standard"

- In Physics, **The Standard Model** is the name given to the current theory of fundamental particles and their interactions

- The Standard Model is a good theory because it has been validated
  - Its predictions have matched experimental data, decimal place for decimal place, with amazing precision
  - All the particles predicted by this theory have been found

- Can such a theoretical model be found for social systems?
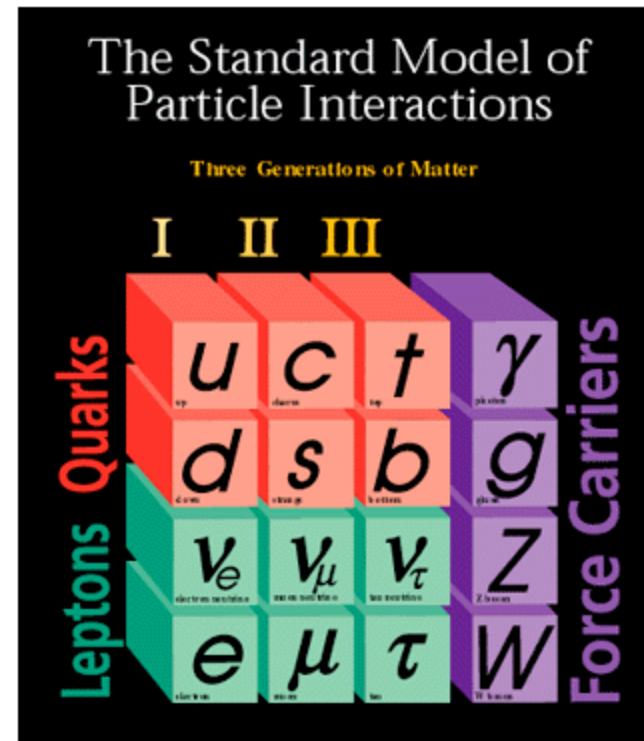
Important Questions about Model Validation

**How can the model be validated if…**

- **Controlled experiments cannot be performed on the system, for example, if only a single historical data set exists?**

- **The real-world system being modeled does not exist?**

- **The model is not deterministic (has random elements)?**

- **How can agent-based models be validated?**
  - **Agent behaviors and interaction mechanisms**
  - **Adaptive agent behaviors of emergent organizations**

Model Validation:
Does the model represent and correctly reproduce the behaviors of the real world system?

- **Validation ensures that the model meets its intended requirements in terms of the methods employed and the results obtained**
- **The ultimate goal of model *validation* is to make the model useful in the sense that the model addresses the right problem, provides accurate information about the system being modeled, and to makes the model actually used**
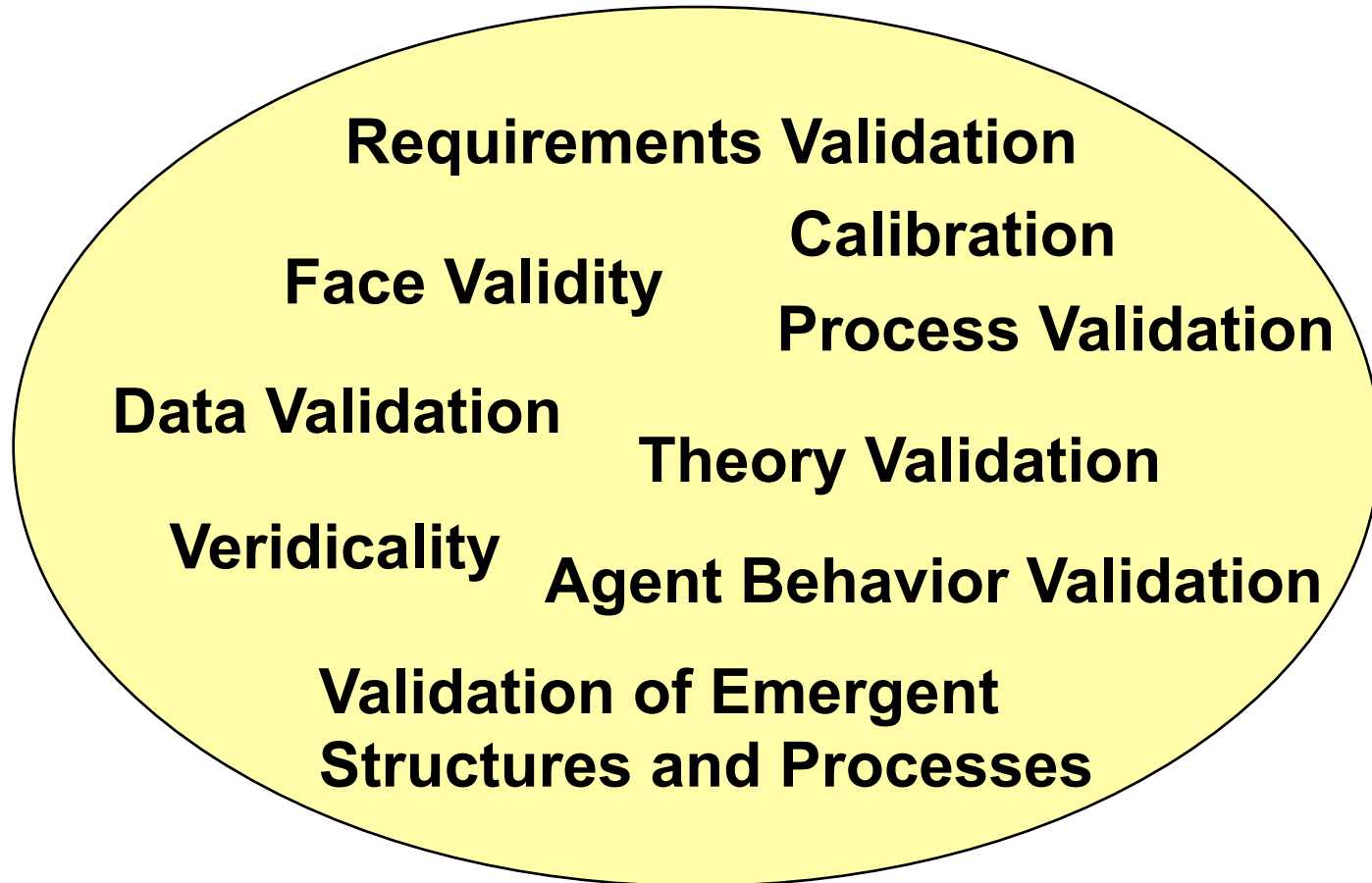
An Information Focused View of Model V&V
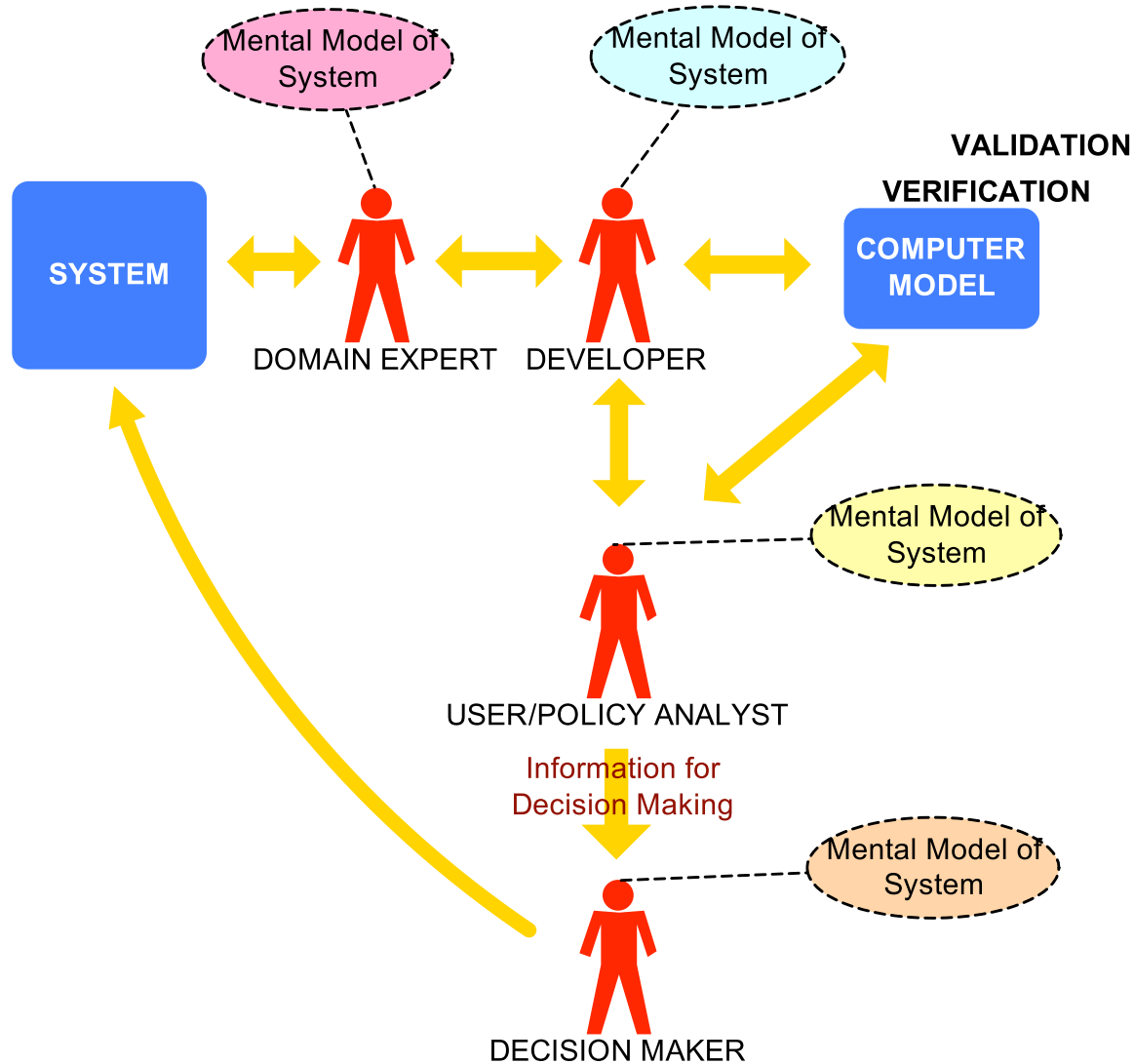
Validation Depends on the Purpose of the Model and Its Intended Use

- **Reasons we do modeling and simulation:**
  - We are constrained by linear thinking: We cannot understand how all the various parts of the system interact and add up to the whole
  - We cannot imagine all the possibilities that the real system could exhibit
  - We cannot foresee the full effects of cascading events with our limited mental models
  - We cannot foresee novel events that our mental models cannot even imagine
- **We model for insights, not numbers**
  - As an exercise in "thought space" to gain insights into key variables and their causes and effects
  - To construct reasonable arguments as to why events can or cannot occur based on the model
- **We model to make qualitative or quantitative predictions about the future**

The World of Model Validation

Requirements Validation

Calibration

Face Validity

Process Validation

Data Validation

Theory Validation

Veridicality

Agent Behavior Validation

Validation of Emergent Structures and Processes

Practical Validation

- **Validation exercises amount to a series of attempts to invalidate a model**
  - One recently proposed V&V technique, Active Nonlinear Tests (ANTs), explicitly formulates a series of mathematical tests designed to "break the model"
- **Presumably, once a model is shown to be invalid, the model is salvageable with further work and results in a model having a higher degree of credibility and confidence**
- **The end result of validation**
  - Technically not a validated model, but rather a model that has passed all the validation tests
  - A better understanding of the model's capabilities, limitations, and appropriateness for addressing a range of important questions

Establishing Credibility

- Unlike physical systems, for which there are well-established procedures for model validation, no such guidelines exist for social modeling
- In the case of models that contain elements of human decision making, validation becomes a matter of establishing *credibility* in the model
- Verification and validation work together by removing barriers and objections to model use
- The task is 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
- The process is akin to developing a legal case in which a preponderance of evidence is compiled about why the model is a valid one for its purported use

Pathways to Validation

- **Cases**
    - **Exploration of critical cases**
    - **Exhaustive exploration of cases**
- **Using models as exploratory e-laboratories**
    - **Rapid prototyping**
- **Multiple models**
- **Maximally diverse model ensembles**
- **Using subject matter experts**
    - **Evaluation**
    - **Role playing, participatory simulation**
- **Computational simulations as a special cases of analytical modeling**

Validation Case Study:
Deregulated Electric Power Markets

- **Data**
  - Checking the currency of the data with the original data sources
  - Cross-checking data with third parties having a vested interest in the data
- **Subject Matter Experts (SME)**
  - Model was developed by a team of experienced domain experts
  - Independent electric utility SMEs provided critical industry experience
- **Participatory Simulation**
  - Ability to place themselves in the positions of agents in the deregulated markets
- **Replication of Special Cases**
  - Model runs constructed to replicate special case for the previously validated *regulated* power market
- **Comprehensive Model Cases for the Agent Parameter and Strategy Space**
  - Not possible to draw general conclusions from only a handful of model runs: non-linear, dynamic aspects of the agent behaviors and interactions
  - Extensive cases verified expected model behaviors and discovered model unexpected model behaviors
  - Unexpected cases created focal points for further model runs and in-depth analysis
  - Comprehensive testing of plausible agent strategies
  - Extensive use of data visualization techniques
- **Model-to-Model Validation**
  - Validation of the simplified DC model to the complete and validated AC model was done by comparing results for extensive number of cases

Validation Case Study: Deregulated Electric Power Markets

- **Lessons**
  - **All model results and the answers to the obvious questions pertaining to the model results had to be explainable in plain English or they would not be useful to decision makers**
  - **The model validation phase ended up taking as long as the model development phase. In the end, however, it was generally accepted that the model was a valid one for answering a wide range of important questions pertaining to electric power deregulation**

Example Supply Chain Results Over Fine Mesh

Total Cost
$\theta = 0.25, Q = 17$

Agent Behavioral Parameters
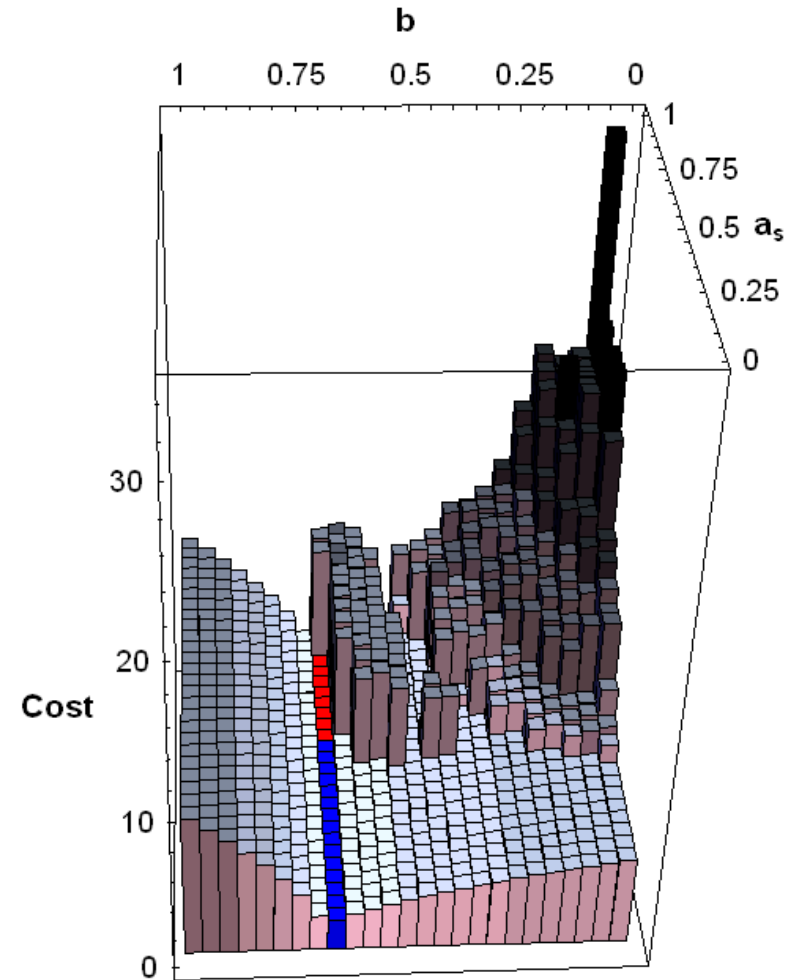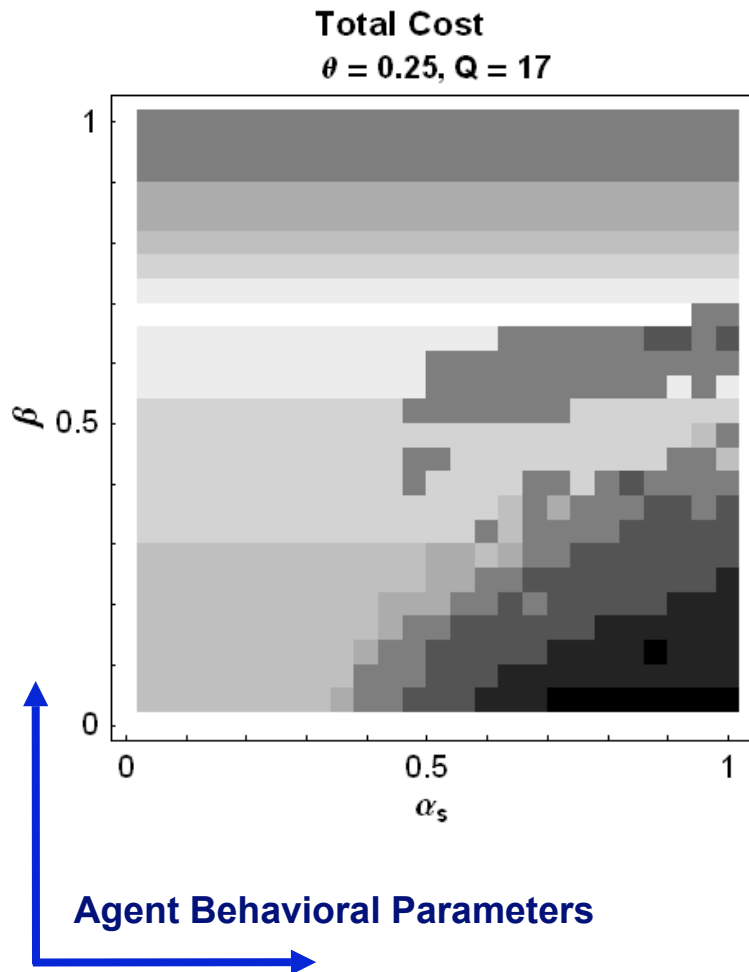
Other Aspects of V&V

- *Independent Verification & Validation*: The V&V of a model is considered *independent* (IV&V) when it is conducted by knowledgeable people other than the original model developers.
- *Accreditation* (IVVA) is the process of determining whether a model is useful for a particular purpose and is applicable to answering a specific set of questions.
- *Certification* is the process of ensuring that a model meets some specified standard or set of standards

The Challenge of Validating Theory

- **Theory validation relates to the technical details of the model and how it relates to the relevant disciplines, knowledgeable expertise and underlying theories**

- **V&V is required at multiple scales**
  - **Agent-to-agent interactions**
  - **Organizations**
  - **Society and culture**

- **Validation of theory**
  - **What theory is used in the models**
  - **How the theory is used in the models**
  - **How the theories are combined in the models**

References

- AIAA. 1998. AIAA Guide for the Verification and Validation of Computational Fluid Dynamics Simulations. American Institute of Aeronautics and Astronautics.
- Annino, J. S. and E. C. Russell. 1979. "The Ten Most Frequent Causes of Simulation Analysis Failure – and How to Avoid Them!" *Simulation* 32(6):137-140.
- Axelrod, R. "Advancing the Art of Simulation in the Social Sciences" *Simulating Social Phenomena.* Conte, R., R. Hegselmann and P. Terna, eds. Springer: Berlin.
- Axtell, R., R. Axelrod, J. Epstein and M. D. Cohen. 1996. "Aligning Simulation Models: A Case Study and Results" *Computational and Mathematical Organization Theory* 1:123-141.
- Bankes, S. C. 1993. "Exploratory Modeling for Policy Analysis" *Operations Research* 41(3):435-449.
- Bankes, S. C. 2002. "Tools and Techniques for Developing Policies for Complex and Uncertain Systems" *Proc. National Academy of Sciences* 99(90003):7263-7266 http://www.pnas.org/cgi/content/abstract/99/suppl_3/7263.
- Burton, R. "Validating and Docking: An Overview, Summary, and Challenge" *Simulationg Organizations: Computational Models of Institutions and Groups.* Prietula, M. J., K. M. Carley and L. Gasser, eds. MIT Press: Cambridge, MA.
- Carley, K. M. 2002. "Simulating Society: The Tension between Transparency and Veridicality" Agent 2002 Conference on Social Agents: Ecology, Exchange, and Evolution. Macal, C. M. and D. L. Sallach, eds. Chicago, IL. Oct. 11-12, 2002. pp. 103-113.
- Edmonds, B. and D. Hales. 2003. "Replication, Replication and Replication: Some Hard Lessons from Model Alignment" Journal of Artificial Societies and Social Simulation 6(4) http://jasss.soc.surrey.ac.uk/6/4/11.html.
- Epstein JM, Axtell R. 1996. Growing Artificial Societies: Social Science from the Bottom Up. Cambridge, Mass.: MIT Press
- Gilbert N, Troitzsch KG. 1999. Simulation for the Social Scientist. Buckingham: Open University Press

References (cont'd.)

- Hales, D., J. Rouchier and B. Edmonds. 2003. "Model-to-Model Analysis" Journal of Artificial Societies and Social Simulation 6(4) http://jasss.soc.surrey.ac.uk/6/4/5.html.
- Harmon, S. Y., C. W. D. Hoffman, A. J. Gonzalez, R. Knauf and V. B. Barr. 2002. "Validation of Human Behavior Representations" *Foundations for V&V in the 21st Century Workshop (Foundations '02)*. The Johns Hopkins University Applied Physics Laboratory, Laurel, MD. October 22-24, 2002. pp. 1-34.
- Hodges, J. S. 1991. "Six (or So) Things You Can Do with a Bad Model" *Operations Research* 39(3): 355-365.
- Law, A. M. and D. Kelton, W. 2000. Simulation Modeling and Analysis, 3rd Ed. McGraw-Hill: New York, NY USA, pp. 800.
- Law, A. M. 2004. "Eleven Critical Pitfalls in Simulation Modeling" *available at http://www.averill-law.com/simulation-news-papers.htm*.
- Lempert, R. J., S. W. Popper and S. C. Bankes. 2003. *Shaping the Next One Hundred Years: New Methods for Quantitative Long-Term Policy Analysis.* RAND:Santa Monica, CA, pp. 187.
- Miller, J. H. 1998. "Active Nonlinear Tests (Ants) of Complex Simulation Models" Management Science 44:620-830.
- Robinson, S. 1999. "Three Sources of Simulation Inaccuracy (and How to Overcome Them)" *Proceedings of the 1999 Winter Simulation Conference.* Farrington, P. A., H. B. Nembhard, D. T. Sturrock and G. W. Evans, eds. pp. 1701-1708.
- Sargent, R. G. 2003. "Verification and Validation of Simualtion Models" *2003 Winter Simulation Conference.* Chick, S., P.J.Sanchez, D. Ferrin and D. J. Morrice, eds. pp. 37-48.
- Sterman JD. 1987. Testing Behavioral Simulation Models by Direct Experiment. Management Science 33: 1572-92.