Challenges in the Operational Use and Validation of Sociocultural Techniques, Tools, and Models

Jonathan Pfautz, David Koelle, Brian Prue, & Corey Lofdahl

Presented at the 2015 NDIA Human-Systems Integration Conference
Part of Session 4: Social, Cultural, Behavioral Understanding (SCBU)

10 February 2015
Alexandria, Virginia

charles river analytics

Outline

- Context: Non-Kinetic Operations
- Motivation & History
- Human-System Challenges for Modeling
  - User analyses
  - Technologist's perspective
- Potential Solutions
- Conclusions & Future Research

Context: Non-Kinetic Operations

- Non-Kinetic Operations (NKO) are activities that do not focus on destroying enemy forces through the application of physical effects
  - E.g., informing a population about where to seek medical aid
  - E.g., urging people to stay safely inside during a civil disturbance
- Our objective: Apply methods for understanding socio-cultural and behavioral factors to aid performance in NKO
  - Guide analysis and understanding of human behavior
  - Enable reasoning about current and likely future behaviors
  - Help identify data needs and integrate data sources
- Results: Developed, deployed, and now sustaining software incorporating multiple socio-cultural techniques to 4,000+ users worldwide
  - Fully integrated into training
  - Part of long-term Program of Record

Motivation & History

- Our starting point: Build computational behavior models
  - Develop and verify models
  - Embed in decision aid to provide “answers”
- Why?
  - Encode and share established knowledge from academic disciplines about human behavior
  - Compactly capture complex and rich knowledge
  - Formalize such knowledge and better enable validation
  - Automate some forms of analysis
  - Process more data faster, more cheaply
  - Provide systematicity and rigor in operational application of scientific knowledge

This approach was not tenable… why?

Behavior Modeling is Difficult

- Creating a model requires defining and understanding:
  - Types of behaviors and internal states to model
  - Theories considered (and deconflicted) – and why
  - Computational representations used (and integrated) – and why
  - Assumptions made
  - Time/cost/scope constraints
- Verifying and validating complex socio-cultural and behavior models is similarly difficult
  - Data may be of wrong type or format, sparse, uncertain, noisy
  - Academic community lacks standards for verification and validation
  - Models are difficult to compare

(Zacharias, MacMillan, & Van Hemel, 2008)

- Bootstrapping across models is hard due to lack of standards; models are not easily composed

(Davis & Anderson, 2004)

But even validated models still are always not used… why?

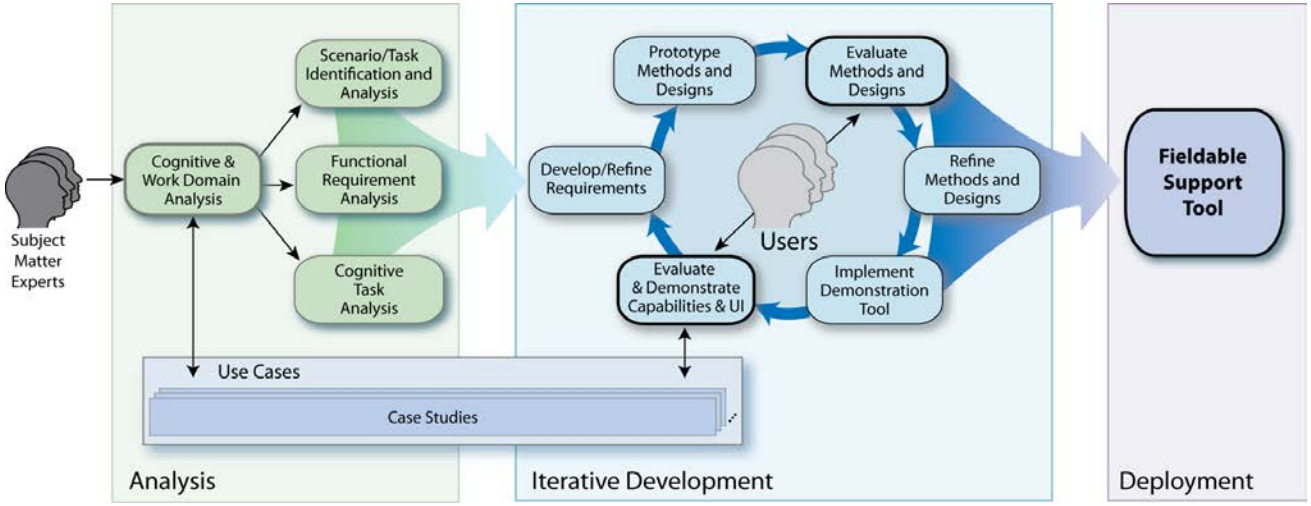
Approach: Analyze the Domain

- How?

Cognitive Task Analysis (Schraagen, Chipman, & Shalin 2000), Cognitive Work Analysis (Vicente, 1999), Work Centered Support Systems (Eggleston, Roth, & Scott 2003), Applied Cognitive Task Analysis (Militello & Hutton 1998), Requirements Analysis, Hierarchical Task Analysis (Shepherd, A. 2000), Goal-Directed Task Analysis (Endsley 1995), …

  - All of these require management of cost/scope trade-off! (Pfautz & Roth 2006)

- Our approach:

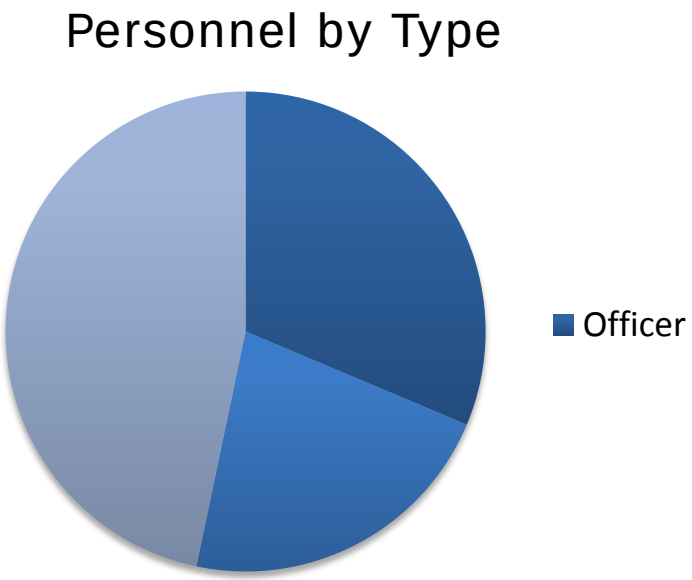


Example Activities
Define cognitive (and other) tasks
Identify challenges and complexities
Develop scenarios to aid analysis and evaluation
Define constraints and opportunities afforded by work environment
Identify socio-organizational considerations
Catalogue existing tools, systems, and data

Analyzing the Domain: Our User Interactions

- 4,000+ total hours of interviews, demonstrations, and evaluations
- 400+ active-duty personnel and Government civilians, spanning:
  - Deployments to >80 different units in garrison, in theater, and during training, command-post, and multiple field exercises
  - Organization/deployment types:
    - Small vs. large groups
    - Varying levels of leadership understanding and accountability
    - Different parent-organization goals
  - Personnel types:
    - Novices and experts
    - Operators, leadership, other managers

We are extraordinarily grateful for our past and ongoing interactions with the user community



Analyzing the Domain: (Some of) Our Results

- NKO tasks are broad and complex
  - Not simply "Will this work?" and "Did that work?"
- Users were skeptical of behavior models at best and dismissive at worst (Farry et al. 2010, Thornton et al. 2010, Carlson et al. 2011)
- Users are highly experienced at operating under uncertainties
  - And require explicit expression of qualifying information (Bisantz et al. 2009)
- Users range significantly in knowledge of the theoretical bases of human behavior
  - E.g., from high school education to PhDs in cultural anthropology
  - Matching user skill/training is essential – in behavior *and* in computer systems
- Usability and utility were paramount
  - Need effective communication of capabilities/assumptions (Pfautz et al. 2009)
  - Logistics matters - "I'm not allowed to install this…" or "Where do I get the data?"

Analyzing the Domain: Observations on Trust in Models

- Trust varies by application
  - Trusting autopilot ≠ trusting a decisions based on adversary behavior model
- Individual and socio-organizational experience with models matters
- Inherent skepticism when source of information is not known, personally
- Inherent skepticism of computational systems
- Skepticism varies as a function of dynamics and criticality of situation

(Farry, Pfautz, et al. 2010)

Observed expressions of model trust:

Use model to provide "truth"
Use model to provide answers, with caveats
Use model to refine own reasoning
Use model to derive additional insights
Use model to confirm own reasoning
Use model to justify reasoning to others
Ignore model completely
Actively disparage model

Note that the level of trust expressed was not always appropriate!

Human-Systems Challenges for Modeling: A Modeler's Perspective

Given the user's perspective, what are the resulting challenges for modelers?

- Access to potential users; unclear user community
  - Avoid assumptions about skills/knowledge across computer and behavioral systems
- Information on typical and current user tasks, mission parameters, and/or specific situational/contextual information
  - Focus on a useful level(s) of analysis, on relevant problems
  - Understand subtleties of model trust for individual & organization
- Access to operational data (and/or types and formats)
- Information on reporting requirements for models
  - May need to show data used, assumptions, internal model processes, and the implications of all of these on the quality of results
- Evaluation of model utility; rethinking notions of "validity"

Human-Systems Challenges for Modeling: Solutions Reconceiving Notions of Validity

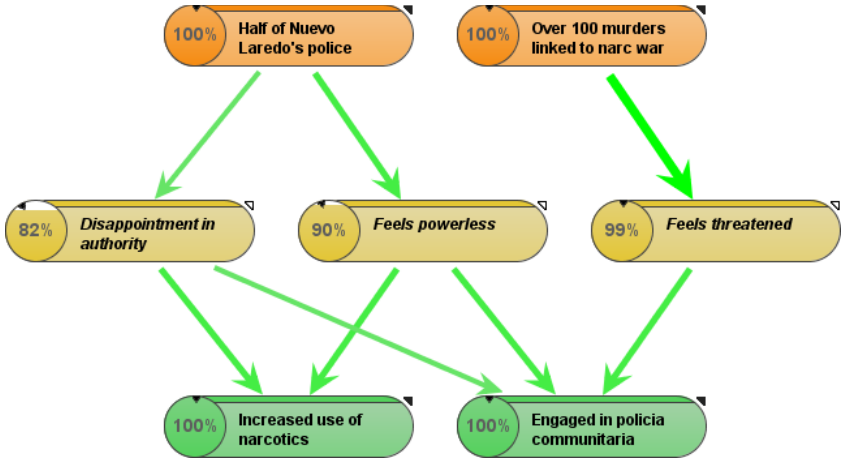
- "A model's capability to serve an applied goal ...is not necessarily equivalent to its construct validity" (Campbell & Bolton, 2005)
- *Application Validity:* The degree to which a model is a faithful representation of the real world from the perspective of the intended use of that model or simulation (DMSO, 2001)
- A human-centered perspective:
  - Bounds the scope of validation
  - Helps understand what tasks, metrics, and methodologies could be used to establish application validity
  - Broadens value of behavior modeling from just "models"
  - Increases likelihood of operational utility and real-world use of models

Human-Systems Challenges for Modeling: Solutions Broadening From HSCB *Models* to HSCB *Methods*

- Individual aspects of the HSCB modeling enterprise may have more value to users than a "complete" model
- Models could:
  - Serve as "references" – a concise communication of knowledge about complex human phenomena, incl. counter-examples (Pfautz et al. 2010)
  - Act as a framework for eliciting expert knowledge to inform analysis and/or decision-making
  - Inform data collection, data fusion, and data interpretation (Mahoney et al. 2011)
  - Act within "meta-models" that help users understand when/where models are applicable (Kettler et al. 2011)
- Modeling formalisms and methodologies could:
  - Streamline expression of situation-specific or general causal knowledge about human behavior (Rosenberg et al. 2011)
  - Improve user's reasoning processes and fact-checking (Cao et al. 2009)
  - Speed up model authoring and validation cycles

Human-Systems Challenges for Modeling: Solutions
User-Created Modeling

- Provide workflow and guidance to lead user in creating model of a population
  - Guidance is derived from social/behavioral methodologies
  - Creates consistency and rigor across users
  - User's reasoning is captured and communicated clearly
  - Vetted sources are integrated to provide audit trails
- Requires formalism(s) that enable rapid user model creation

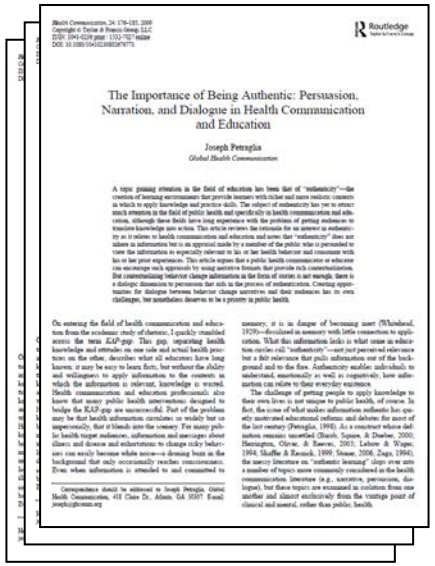


E.g.,

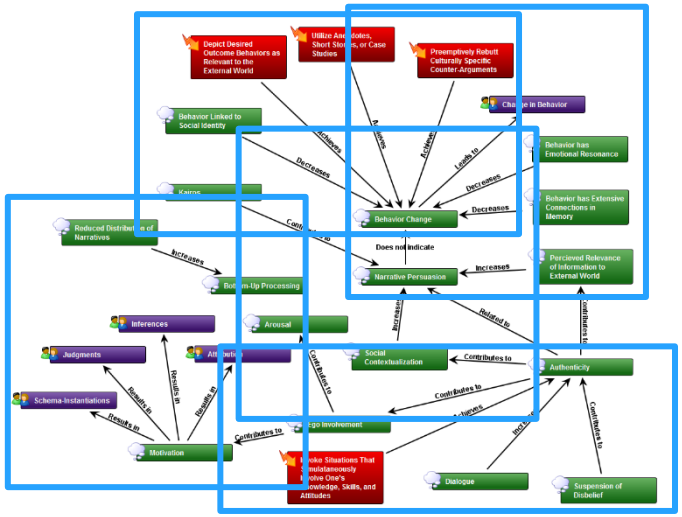
- Causal Influence Models (Pfautz et al. 2010)
- Argumentation systems
- Utility diagrams
- Causal concept maps
- Sensitivity analysis
- Decision trees
- Reference models

Human-Systems Challenges for Modeling: Solutions
User-Adaptable Modeling

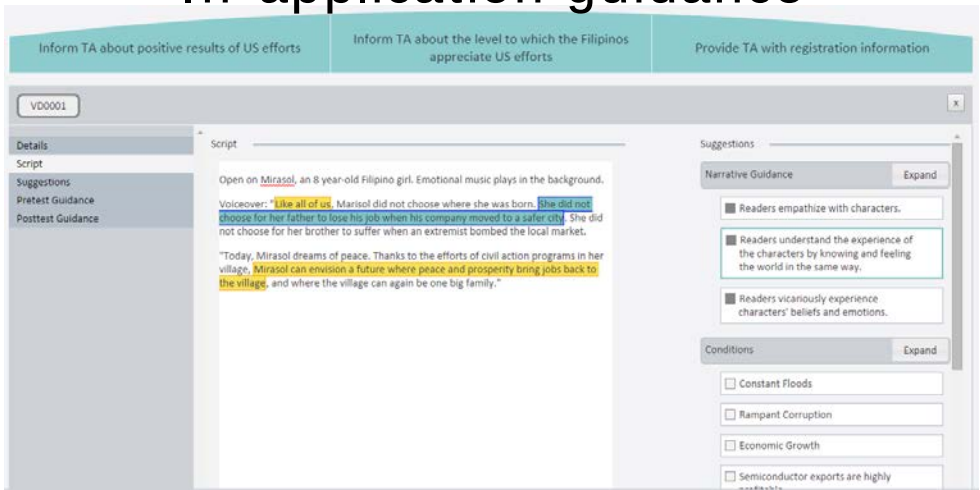
Established Research



User-Selectable "Mini-Models"



In-application guidance



Human-Systems Challenges for Modeling: Solutions Metrics for Evaluation of Operational Utility

- Understanding the domain enables definition of metrics
- Example metrics:

> Task performance (perceptual, cognitive, decision-making, communication), response time, team performance, trust, workload, situation awareness, communication efficiency, psychophysiological correlates, neurological responses, …

- Our research and analysis suggests focusing on:
  - Cognitive/Decision-making task performance
    - On well-defined tasks
    - Across roles/responsibilities within an organization
  - Usability – and its interaction with utility
  - Trust
  - Workload

Wait! What about "Model correctly ingests operational data" and the like?

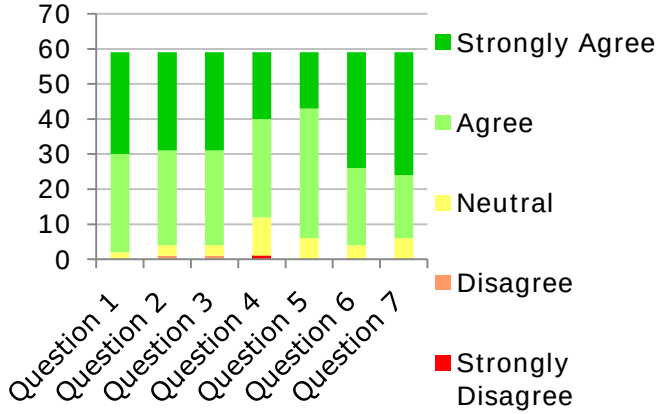
Challenges in Evaluation of Operational Utility

- Motivation of user's leadership
  - Why should I provide subjects for this evaluation?
- Motivation of user population
  - "How do I get on the team that gets to use this? I want to pass."
    – Soldier, during a training culmination exercise
- Access to user population
  - Classified environments and competition against day-job
- Directability of user population
  - May have only secondary or tertiary ability to direct evaluations
- Motivation of sponsor
  - E.g., We receive little support for formal studies of operational utility
    … But are often expected to achieve utility anyway
- Perceptions of the modeling community
  - "User-centered evaluations are just 'the engineering'"

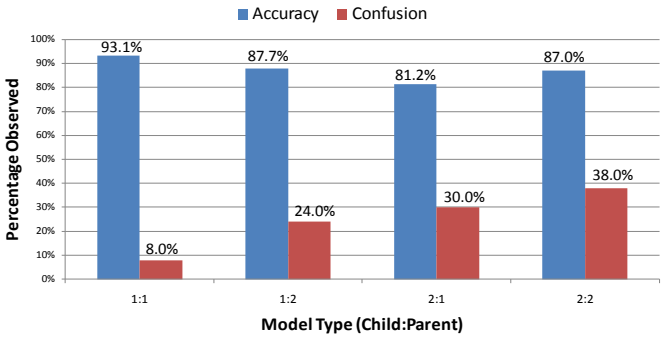
Human-Systems Challenges for Modeling: Solutions Evaluation of Operational Utility – One Approach

- Opportunistic participation in demonstrations and exercises
- Surveys and questionnaires
- Observation and analysis
- Analysis of work products within and without our software
  - By our team
  - By instructors and other experts
- Integration with training assessments
- Use of surrogate populations for formal studies of specific methods and technologies

Q6: I would recommend to my Commander that this tool become a part of our standard software toolkit



U. Buffalo study: Can users represent local knowledge accurately in modeling formalism X?



Conclusions

- Exploiting socio-cultural and behavioral modeling techniques and technology is a bi-directional challenge
  - Need user community engagement
  - Need adaptable S&T community perspectives on modeling – merging operational utility and scientific contributions
- Potential solution space is large and ripe for novel approaches across application domains
- Need ongoing focus on evaluation of both validity and utility
  - Big science-side challenge: “User studies are not relevant to computational social science”
  - Big operations-side challenge: “Prove formal user evaluation is valuable and cost-effective”

Acknowledgments and Questions??

This work presented here includes insights garnered across many efforts conducted under contract from AFRL, OSD, CTTSO, and many others. We also remain grateful for the continuing contributions of individuals conducting non-kinetic operations world-wide.

Dr. Jonathan Pfautz
Senior Principal Scientist

jpfautz@cra.com

Charles River Analytics Inc.
625 Mount Auburn Street
Cambridge, MA 02138
Voice: 617-491-3474
Fax: 617-868-0780
www.cra.com

Bibliography

Bisantz, A., Stone, R., Pfautz, J., Fouse, A., Farry, M., Roth, E., Nagy, A., & Daniels, G. (2009). "Visual Representation of Meta-Information," J. of Cognitive Engineering and Decision Making, 3(1), pp. 67-91.

Campbell, G. & Bolton, A. (2005). HBR Validation: Integrating Lessons Learned From Multiple Academic Disciplines, Applied Communities, and the AMBR Project. In R. Pew & K. Gluck (Eds.), *Modeling Human Behavior With Integrated Cognitive Architectures* (pp. 365-395). Mahweh, NJ: Lawrence Earlbaum Associates.

Cao, C., Guarrera, T., Jenkins, M., Pennathur, P., Bisantz, A., Stone, R., Farry, M., Pfautz, J., & Roth, E.(2009). "Evaluating the Creation and Interpretation of Causal Influence Models," in the *Proc. of the Human Factors and Ergonomics Society 53rd Annual Meeting*, San Antonio, Oct. 19-21.

Carlson, E., Pfautz, J., & Koelle, D. (2010). "Operator Trust in Human Socio-Cultural Behavior Models: The Design of a Tool for Reasoning about Information Propagation," in *Proc. of AHFE*, Miami, July 17-20.

Davis, Paul K. and Robert H. Anderson. 2004. *Improving the Composability of Department of Defense Models and Simulations*. Santa Monica, CA: RAND.

Eggleston, R. G., Roth, E. M., & Scott, R. A. (2003). A Framework for Work-Centered Product Evaluation. In Proc. of Human Factors and Ergonomics Society 47th Annual Meeting. Denver, CO: Human Factors and Ergonomics Society.

Endsley, M. R. (1995). "Measurement of situation awareness in dynamic systems." *Human Factors*, 37 (1), 65-84 3)

Farry, M., Pfautz, J., Carlson, E., & Koelle, D. (2010). "Trust and Reliance in Human Socio-Cultural Behavior Models," in *Proc. of AHFE*, Miami, July 17-20, 2010.

Kettler , B. & Hoffman, M. (2012). "Lessons learned in instability modeling, forecasting and mitigation from the DARPA Integrated Crisis Early Warning System (ICEWS) Program," in *Proc. AHFE 2012,* July 22-25, San Francisco, CA.

Lofdahl, C., Pfautz, J., Farry, M., & Stickgold, E. (2014). Identifying and understanding trust relationships in social media. In C. Ehlschlaeger (Ed.) *Socio-Cultural Analysis with the Reconnaissance, Surveillance, and Intelligence Paradigm.*

MacMillan, J., Zacharias, G., Freeman, J., and Bullock, B. (2011). "Strategic Roadmap for Human Social, Cultural, and Behavioral Science and Technology," *Proc. of Human Social Culture Behavior (HSCB) Focus 2011,* Chantilly, VA.

Mahoney, S., J. Pfautz, J., Campolongo, J., and Thomsen, E. (2011). "A Semi-Automated Interpretation Tool for Socio-Cultural Data Collection," Proc. of Human Social Culture Behavior (HSCB) Focus 2011, Chantilly, VA.

Militello, L. G. & Hutton, R. J. B. (1998). Applied Cognitive Task Analysis (ACTA): A Practitioner's Toolkit for Understanding Cognitive Task Demands. *Ergonomics*, 411618-1641.

Pfautz, J., Carlson, E., Farry, M., & Koelle, D. (2009a). "Enabling Operator/Analyst Trust in Complex Human Socio-Cultural Behavior Models, " in the *Proc. of the Human Behavior-Computational Intelligence Modeling Conference 2009*, Oak Ridge, Tennessee, June 23-24.

Pfautz, J., Carlson, E., Koelle, D., & Roth, E. (2009b). "User-Created and User-Adaptable Technosocial Modeling Methods," in the *Proc. of the AAAI Spring Symposium on Technosocial Predictive Analysis.* Stanford University, March 23-25.

Pfautz, J., Carlson, E., Koelle, D., & Catto, G. (2010). "Integrating User-Creating Modeling Tools with Complex HSCB Models," in the *Proc. Of 78th Military Operations Research Society Symposium (MORSS)*, June 23rd, Quantico, VA.

Pfautz, J. & Koelle, D. (2014). Approaches to addressing challenges in the operational use and validation of socio-cultural techniques, tools, and models. In C. Ehlschlaeger (Ed.) *Socio-Cultural Analysis with the Reconnaissance, Surveillance, and Intelligence Paradigm.*

Pfautz, J. & Roth, E. (2006). "Using Cognitive Engineering for System Design and Evaluation: A Visualization Aid for Stability and Support Operations," *International J. of Industrial Ergonomics*, 36(5), 389-407.

Rosenberg, B., Furtak, M., Guarino, S., Harper, K., Metzger, M., Neal Reilly, S., Niehaus, J., and Weyhrauch, P. (2011). "Easing Behavior Authoring of Intelligent Entities for Training," in *Proc. of the 20th Behavior Representation in Modeling & Simulation (BRIMS) Conference*. Sundance, Utah. 2011.

Schraagen, J. M. C., Chipman, S. F., & Shalin, V. L. (2000). Introduction to Cognitive Task Analysis. In J. M. C. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive Task Analysis* (pp. 3-24). Mahwah, NJ: Lawrence Erlbaum Associates.

Shepherd, A. (2000). *Hierarchical Task Analysis*., CRC.

Thornton, W., Carlson, E., and Pfautz, J. (2011). "Adapting Social Science To Influence Operations Planners (...And Not The Other Way Around)," Proc. of Human Social Culture Behavior (HSCB) Focus 2011, Chantilly, VA.

U.S. Department of Defense (2001). VV&A Recommended Practices Guide Glossary. Washington, D.C.: Defense Modeling and Simulation Office.

Vicente, K. J. (1999). *Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum Associates.

Zacharias, G. L., MacMillan, J., & Van Hemel, S. B. (2008). *Behavioral Modeling and Simulation: From Individuals to Societies*. Washington, D.C.: National Academies P.