



HSCB

Day 1

Working Group 4

**VV&A of Methods, Tools, & Data
Report**



Working Group 4: VV&A of Methods, Tools, & Data Structure of the Working Group



- Chair: Dr. Dean S. Hartley III
- Co-chair: Dr. Marina Arbetman-Rabinowitz
- Members:
 - Michael Bailey
 - Pauline Baker
 - Chuck Barry
 - Todd Brethauer
 - Alok Chaturvedi
 - Claudio Cioffi-Revilla
 - Rick Cunningham
 - Jack Goldstone
 - Ari Greenberg
 - J.C. Herz
 - Charles Macal
 - Brett Marvin
 - Mary McDonald
 - Jimmie McEver
 - Anne McGee
 - Lisa Moya
 - Jennifer O'Connor
 - Maureen O'Mara
 - Michael Ottenberg
 - CDR Brett Pierson
 - Eunice Santos
 - Gary W. Schaeff
 - Stephen Shellman
 - V.S. Subrahmanian
 - Ben Wise
 - Jordan Willcox
 - Ted Woodcock



Working Group 4: VV&A of Methods, Tools, & Data

Process Description (1)



Key issue: VV&A of models, methods, data, theory

- **Presentations**

- VV&A of DIME/PMESII models: it can be done and how to do it
- Tests and forecasts using theories and models is often inadequate due to problems of validity and reliability of concepts and data. Mapping of solutions.

- **Brainstorming**

- Given the scope and breadth and varying depths of the methodologies (theories and computational modeling techniques), tools (models and systems of models), data related issues and econometric techniques (time series, static, real-time)
 - How should we organize these dimensions to make them comprehensible in a useful way?
 - Assuming VV&A can help us with this organization by structuring what we know about them, how do we make that happen?
 - What do we want from HSCB models, annotated by “wishful thinking,” “possible in the future,” “possible in the near future,” and “possible now”?



Working Group 4: VV&A of Methods, Tools, & Data

Process Description (2)



Brainstorming (Continued)

- Data Requirements – Has this been a limiting factor? Rank 1 (most important limiting factor) -10 (not a limiting factor)
 - Responses 1-2:
 - Responses 3-4:
 - Responses 5-6:
 - Responses 7-8:
 - Responses 9-10:
- How much time is allocated in your projects to data related items?
 - Responses 0-10%:
 - Responses 11-20%:
 - Responses 21-30%:
 - Responses 31-40%:
 - Responses 41-50%:
- How can data help improve your efforts? What are the gaps? Ideal world?
- Note: There was not enough time to discuss these questions during the workshop. Questions were emailed to the participants. The answers will be collated.



Organizing principles

for Theories, Tools, Algorithms, Datasets (First Brainstorming Question)

- **Tags (partial listing)**
 - **General characteristics**
 - **Systemic Organization – Inputs, Processes, Outputs**
 - **Academic Discipline**
 - **Function – Outcomes produced**
 - **Assumptions Used**
 - **Level (strategic, operational, tactical)**
 - **Domains: land, sea, air, space, cyberspace**
 - **Time – Forecasting, Assessment, Historical**
 - **Granularity: Sub-individual, Individual, Group,**
 - **Coverage of DIME/PMESII**
 - **Methodological Approach**
 - **Representation: Descriptive, Observation, Intervention**
 - **Secondary characteristics**
 - **Update rates on Data**
 - **Update to the Model (granularity in time)**
 - **Military Functions, maneuver, fires, C2, intel, sustainment**
 - **Surfboards: training, analysis, T & E, planning, operations**
 - **# of sides – where can model be used ..**
 - **Deterministic vs. stochastic, closed form vs. human in the loop.**
 - **Pedigree- metadata data – characteristics of data**
 - **User Needs characteristics**
 - **Owner / How do I acquire it (COTS, MOTS, web services)**
 - **Ease of use**
 - **What is time / effort / cost of usage**
 - **Maintenance (software & training) requirements**
 - **Interoperability with other systems (DoD, agencies, allies)**
 - **Tactically mobile, reliable, networked**
- **Visualization add-ons : operational use of tags; requires standard tags**
- **V&V to be useful ought to help categorize model data in this sense: way to create tags**



Organizing principles for Empirical Work

Often the Failure of Predictive Models is due to inaccurate specification and poor data collection:

- Fundamentals of applied work require a **problem oriented approach** rather than technique oriented approach
- **Hierarchical Modeling:** The organizing question should navigate across theories, models and levels of analysis
- **Specifications of variables:** require Concepts and Operationalization (semantics and ontology)
- **Inventory** of concepts, variables and specifications is needed
- Data Integrity, Consistency, Maintenance, and Dissemination are required
- **Reliability** and **Validity** Tests of operational qualitative and quantitative variables must be performed



What should a model do for us?



A model should answer questions (Third Brainstorming Question)

- Why people become insurgents?
 - Why are people X becoming Insurgent Y in location Z? •<=2 years
 - Which of these people are becoming insurgents? –~5 years
 - How do I “reverse this”? How long? What will it take? –>5 years
 - How are the red team going to respond to our actions? Who is the red team? –~5 years
- Given limited \$\$, what part of DIME should we invested in? –?
- Given desire for success, which COA should be done? •~5 years
- Tools that provide situational awareness and short-term forecasting. (Desired) •?
- What are the consequences including unintended of our actions? •~5 years
- How robust are results against uncertainty? Model incompleteness – other possible answers. •>5 years
- What are the system drivers / interaction mechanisms of the results. •~5 years
- Scalable organizational performance model of social institutions [city/town]. (MPICE). •~5 years
- Model that accommodates all theories, or knows why not. •>5 years
- Bridge training/education/operations – all tailored to be the same, see book “Ender’s Game”. •>5 years
- Do we want a GIG (global information grid) service of DIME/PMESII. We want something that plugs in a question and spits out an answer relative to the population. •~5 years
- Technical –
 - models should measure outcomes and not inputs. Not by what you do, but by what happens when you do it. –?
 - Standardized, interoperable data. Needs to be available on point to the end user and automatically enforced. Knowledge of what model does not do. –<=2 years



Working Group 4: VV&A of Methods, Tools, & Data

Results (1)



- VV&A issues
 - Definitions
 - Verification is the process of determining that a model or simulation implementation accurately represents the developer's conceptual description and specification
 - Validation is the process of determining the degree to which a model or simulation is an accurate representation of the real-world from the perspective of the intended uses of the model or simulation.
 - Accreditation is an official determination that a model is acceptable for a specific purpose
 - Actually doing VV&A
 - Standards of accomplishment
 - Agent-based models
 - Theology: are precise definitions, levels of accomplishment, descriptions of pitfalls more important than getting started with performing VV&A?
- Data VV&A issues
 - DIME/PMESII data V&V has similarities to combat model data V&V; however, significantly larger complexity of the data sets means the V&V is done at the variable level by [time span & country (or unit of analysis), conceptual definition]
 - [Standard/original source] Data sets are never really ready for use – require clean up to fit current need
 - Need Central Library with concepts, variables, metadata (Data dictionaries, semantic descriptions, ontologies), granularity & validation, needs to include qualitative data as well as quantitative data
 - A concepts to measurements workshop would be useful
 - Outcomes vs. outputs [counting # of flyers distributed is not same as measuring the impact of the information operation]
- Theories & methods need V&V equivalent, just as have need for VV&A of data & tools
 - In the sense of detailed “examination” of same to produce “tags”
 - Need standards for the tags



Working Group 4: VV&A of Methods, Tools, & Data

Results (2)



- Lack of data
 - Wealth of models
 - No data or insufficiently “good” data to test them
 - How to look at what is available, what is not available –
 - create synthetic data? Anonymous data?
 - Who can get best data – not necessarily the classified data
 - Database maintenance & poor dissemination of data & metadata
 - Models that use “similar” data to the data that are really needed for the model may have dissimilar outcomes from what they ought to have – because there is an additional implicit model of the real data by the “similar” data
 - This is a thing that VV&A needs to look for
- Legal issues
 - Data Ownership
 - Who owns data/intellectual property
 - If the government owns it, it may be subject to FOIA release
 - It may be illegal for the government to own or even store some data
 - Government sponsored projects and attendant data restrictions and requirements
 - Consequences
 - What happens when data collected for one purpose (passive) is made available to support US military operations? Are original (non-military) researchers potentially liable?
 - What happens when incorrect data is used to formulate policy that harms innocent individuals?
 - Lawyers to keep you from being sued when someone you interviewed is killed
 - FISA and the Telecom companies
 - Distribution restrictions
 - Classification
 - Academic retention until publication
 - Privacy issues



Working Group 4: VV&A of Methods, Tools, & Data

Results (3)



- The want is for data vs. model co-evolution
- Deciding how close we are to being able to do X (answering a particular question) is hard (see slide 7)
 - We can almost always build a model that purports to do X
 - How do we know if it really does X?
 - How do we know when we will be able to do it in future?
 - What data are needed to do testing of whether X is done? Time & money are involved in getting the data.
- Added Value of Workshop: Opening of Communications
 - People come up afterwards and say – yes (or no)
 - Working Group is sending out to group for post-workshop feedback: Group could not agree on whether these some **could** be answered.



Working Group 4: VV&A of Methods, Tools, & Data

Recommendations to NDU (1)



- Define Data
 - Start with PITF data gathering project and extant data list
 - Consider whether HSCB modeling requires a larger data set
 - Define metadata requirements
 - Parcel out conceptual islands for validation of concepts-measurements
 - Gather and analyze data
 - Commission searches through social sciences professional societies
 - Hold a conference to refine metadata and data V&V
 - Commission a multi-disciplinary group to perform rough V&V for metadata tagging purposes
- Define Theories
 - Start with extant theory list
 - Define metadata requirements
 - Gather and analyze theories to models
 - Commission searches through social sciences professional societies
 - Hold a conference to refine metadata and theory V&V
 - Commission a multi-disciplinary group to perform rough V&V for metadata tagging purposes
- Define Methods
 - Start with the *Encyclopedia of Operations Research and Management Science*
 - Define metadata requirements
 - Gather and analyze methods
 - Commission searches through INFORMS and MORS
 - Hold a conference to refine metadata and method V&V
 - Commission a multi-disciplinary group to perform rough V&V for metadata tagging purposes



Working Group 4: VV&A of Methods, Tools, & Data

Recommendations to NDU (2)



- Define Models
 - Start with the extant list of DIME/PMESII models (Hartley, NRL, JFCOM, OSD PA&E)
 - Define metadata requirements (start with DIME/PMESII Model VV&A Tool and above sources)
 - Commission a multi-disciplinary group to perform rough V&V for metadata tagging purposes
- Hold a Data Library Concept Workshop (after above work)
 - Consider the potential problems of a Data Library
 - Consider the options: central library, virtual library (distributed), government vs academic vs joint
 - Consider the legal issues
 - Consider the data maintenance, creation, VV&A, tagging, and funding issues
- Hold a Measures of Merit (MOMs) Workshop (after first four items)
 - Outcomes vs outputs vs inputs measures [counting # of flyers distributed is not same as measuring the impact of the information operation]
 - Desired MOMs vs availability of data vs model conversion of data to MOMs