



Calhoun: The NPS Institutional Archive
DSpace Repository

Faculty and Researchers

Faculty and Researchers' Publications

2010

Developing Best Practices for Validation of Irregular Warfare Models

Brown, Richard F.; Applegate, Jeffrey; Perkins, Timothy K.;
Blais, Curtis; Duong, Deborah

<https://hdl.handle.net/10945/30785>

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

Developing Best Practices for Validation of Irregular Warfare Models

Captain Richard F. Brown
TRADOC Analysis Center - Monterey
700 Dyer Road, Naval Postgraduate School
Monterey CA, USA 93943
richard.f.brown1@us.army.mil

Timothy K. Perkins
U.S. Army Engineer Research and Development Center
700 Dyer Road, Naval Postgraduate School
Monterey CA, USA 93943
timothy.k.perkins@us.army.mil

Dr. Jeffrey Appleget
Operations Research Department
Naval Postgraduate School
1 University Circle
Monterey CA, USA 93943
jaappleget@nps.edu

Curtis Blais
MOVES Institute
Naval Postgraduate School
1 University Circle
Monterey CA, USA 93943
clblais@nps.edu

Dr. Deborah Duong
Augustine Consulting, Inc.
dvduong@nps.edu

Keywords:

Irregular warfare; validation; cultural geography; methodology; best practices; human behavior modeling;

ABSTRACT: *Operational, analytic, training, testing, and experimentation communities require irregular warfare (IW) methods, models, and tools (MMT) with tactical-level representation to address key issues and requirements. DoDI 5000.61 requires all models and simulations developed, used, or managed by the DoD to be verified, validated, and accredited. Critical to the tactical representation of IW are interactions between combatants and the indigenous population. Representation of these interactions between humans (human behavior representation, or HBR) will require expertise from several of the many fields of social science, and the verification, validation, and accreditation (VV&A) of these representations will require adaptation and, in some cases, enhancement of VV&A techniques that have become common practice when conducting VV&A of DoD physics-based combat models and simulations. Moreover, new techniques may be required. The U.S. Army Training and Doctrine Command (TRADOC) Analysis Center, Monterey (TRAC-Monterey) and the Naval Postgraduate School (NPS) are addressing this need by developing best practices for validation of IW MMT representation and applying those best practices to selected existing IW MMT, namely the Cultural Geography (CG) Model and the Peace Support Operations Model (PSOM). The specific aspect of IW that this effort has focused on is the representation of counterinsurgency environments and operations at the tactical level.*

1. Introduction

The U.S. Army Training and Doctrine Command Analysis Center (TRAC) is currently engaged in a three year campaign plan focused on building its Irregular Warfare (IW) analytical capability. An essential element of this plan is building the proper methods, models, and tools (MMT) that facilitate analyses of IW. Like any MMT used in the analytic community, those that support analyses of IW are subject to the same validation, verification, and accreditation (VV&A) standards as any that support analyses of major combat operations (MCO).

While IW MMT will invariably represent some aspect of combat, they also will represent the center of gravity in the IW environment: the indigenous population. Thusly, IW MMT contains some form of human behavior representation (HBR), and the VV&A standards for such representation are much more difficult to ascertain than the aforementioned MCO type MMT. This paper contains excerpts from the 2010 Best Practices Guide (BPG) v.1.0 prepared by TRAC – Monterey and the Naval Postgraduate School (NPS). The BPG introduces a generic framework toward validation of IW models and this paper provides an overview of two sections of the

BPG. The first section discusses IW, why it is important to the analytic community and some of the challenges inherent to conducting IW analysis. The second section includes the validation framework and also addresses the challenges identified in the first section.

2. Irregular Warfare Modeling

Irregular warfare (IW) is defined as “a violent struggle among state and non-state actors for legitimacy and influence over the relevant populations. IW favors indirect and asymmetric approaches, though it may employ the full range of military and other capabilities, in order to erode an adversary’s power, influence, and will. It is inherently a protracted struggle that will test the resolve of our Nation and our strategic partners” [1]. The crux of this definition is the focus on influence over populations, as contrasted with the goal of major combat operations which is to defeat an opposing military force. In December 2008, DoD Directive 3000.07 elevated Irregular Warfare to be as strategically important as traditional warfare [2].

Irregular Warfare encompasses several types of warfare, from unconventional warfare, to information operations, and even to law enforcement activities that focus on countering irregular adversaries [1]. Counterinsurgency (COIN) operations is one of the focuses for current operations in both Iraq and Afghanistan and is also a type of IW. Essential to a successful counterinsurgency campaign is isolation of the insurgents from the civilian population. FM 3-24/MCWP 3-33.5, Counterinsurgency, states “It is easier to separate an insurgency from its resources and let it die than to kill every insurgent. Clearly, killing or capturing insurgents will be necessary, especially when an insurgency is based in religious or ideological extremism. However, killing every insurgent is normally impossible. Attempting to do so can also be counterproductive in some cases; it risks generating popular resentment, creating martyrs that motivate new recruits, and producing cycles of revenge” [3]. The risks illustrated in this statement stem from the insurgent force’s ability to embed themselves within the population. Thus, separating insurgents from the civilian population, either by physical means or by psychological means, becomes a critical goal of conducting successful COIN operations. In addition, understanding a civilian population’s willingness to support either the government or its opposition (insurgency) becomes a critical measure in the attainment of the goal of isolating insurgents.

While IW is not a new phenomenon, it has typically been the purview of Special Operations Forces (SOF). Since the end of the Vietnam era and until the conflicts in Iraq and Afghanistan, US involvement in IW operations has been on a much smaller scale. The size of these conflicts

has made it necessary to expand the skills of General Purpose Forces (GPF) to include conducting IW, as there is not enough SOF to conduct two, large-scale IW operations simultaneously. Preparing GPF to conduct IW, and preparing US Armed Forces for future operations that will certainly include additional IW challenges, points the DoD Modeling and Simulation community toward the development of Irregular Warfare modeling capabilities to address the operational, analytic, training, testing, and experimentation communities’ ability to represent larger-scale Irregular Warfare operations.

2.1 IW Modeling

Because IW is focused on influencing relevant populations, the focus of IW modeling is substantially different than most existing combat models that represent conflict between two organized, armed (and typically mechanized) forces.

“We lack, as a community, a robust capability to represent, account for, and analyze the Irregular Warfare (IW) environment across the range of tactical, operational, and strategic levels of warfare; as a result we cannot effectively inform decisions concerning operations within the IW environment.”

This statement, taken from a recent briefing to senior DoD leadership, describes the current state of DoD IW modeling efforts today. There are some promising models and modeling efforts that exist or are in development, and there are a few DoD organizations that have begun concerted efforts to understand and clarify the IW M&S challenges. A complete understanding of IW modeling requirements, from the tactical to strategic levels for each of the M&S communities of interest, is just beginning to emerge. However, the critical IW modeling challenges are fairly well known, as they parallel the challenges US forces have encountered in the current IW campaigns in Iraq and Afghanistan.

Key representational challenges facing developers of IW models are described below.

2.1.1 Social Sciences/Human Behavior

Because IW models must focus on representing influence on civilian populations, understanding how to represent and what influences human behavior is critical. This includes both individual and group behavior, understanding the effect of culture on the ability of people to be influenced, and understanding multi-cultural societies. This emphasizes the first substantial difference between combat models and IW models—IW modeling must be informed by the relevant social science

discipline(s). The chart included here (Figure 1) is attributed to Yuna Wong, MCCDC, (Wong, 2009) and although the categorization and strata might be debatable, it is clear that IW model development will require the engagement of social science subject matter experts (SMEs) from many disciplines.



Figure 1. Levels of Analysis in Military OR and Social Science (Wong, 2009)

Because of the unique nature of Irregular Warfare, representation of specific social phenomena has become critical. However, modeling of social science phenomena, while a growing field, is still in its infancy as well. Epstein and Axtell, in the introduction of their book *Growing Artificial Societies* talk about the key reasons social sciences should be considered the “hard” sciences (the following bulleted text is verbatim from the book’s introduction (pp. 1-2); further detailed citation of material is contained in the original text): [4]

- Many crucially important social processes are complex. They are not neatly decomposable into separate subprocesses-economic, demographic, cultural, spatial-whose isolated analyses can be aggregated to give an adequate analysis of the social process as a whole. There is no natural methodology for studying these processes together, as they co-evolve.
- It is difficult to test hypotheses concerning the relationship of individual behaviors to macroscopic regularities, hypotheses of the form: If individuals behave in thus and such a way-that is, follow certain specific rules-then society as a whole will exhibit some particular property.
- The rational actor-a perfectly informed individual with infinite computing capacity who maximizes a fixed (non-evolving) exogenous utility function bears little relation to a human being. Yet, there has been no natural methodology for relaxing these assumptions about the individual.
- It is standard practice in the social sciences to suppress real-world agent heterogeneity in model-building. This is done either explicitly, as in representative agent models in macroeconomics, or implicitly, as when

highly aggregate models are used to represent social processes. There has been no natural methodology for systematically studying highly heterogeneous populations.

- Social science, especially game theory and general equilibrium theory, has been preoccupied with static equilibria, and has essentially ignored time dynamics. There has been no natural methodology for studying non-equilibrium dynamics in social systems.

The five bullets above point to several of the key challenges DoD IW modelers face; namely:

- We need to have models that account for Diplomatic, Information, Military, and Economic (DIME) actions and Political, Military, Economic, Social, Information, and Infrastructure (PMESII) effects.
- We need to understand how human interactions at a tactical level affect the operational and strategic levels.
- The rational actor model, a staple of social sciences that purports to describe human decision making, would have more utility if relaxation techniques to its assumptions existed. In particular, the concept of ‘rational’ will likely vary by culture; e.g. Westerners have a difficult time accepting that suicide bombers are rational actors.
- Modeling civilian populations and their heterogeneity (e.g. some support the government, some support the insurgents, some are neutral) is a key challenge that IW modeling must address.
- The usefulness of equilibrium modeling is unclear at this time; but it is clear that our models need to be dynamic over time. One of the key requirements for IW modeling is representing the impact of actions designed to permanently change the attitudes and behaviors of civilian populations over time.

Fortunately, the state of the art of social science modeling has advanced since Epstein and Axtell’s 1996 assessment above. That said, many of the challenges they identified still remain, in some form, today.

Thus, the challenge for DoD IW modeling is not solved by simply finding the social science expertise and then tapping existing knowledge, models, and data—DoD must engage the best and brightest of the social science community and, as a team, develop solutions to challenges that both communities understand are necessary and difficult.

2.1.2 Non-aggregation/Complex Behaviors

There are two critical distinctions between social science modeling and the physics-based modeling that underpin DoD's combat M&S. First is the stratification of the levels of war. DoD combat models typically fall into strategic, operational, or tactical levels of representations, a separation that is both logical, and fits neatly within doctrine. There is no such clear distinction for IW representation. Tactical actions such as presence patrols can be conducted for months with no apparent effect at any level until the one patrol that detains a key militia leader, and then the effect of the capture ripples throughout the country. "The Strategic Corporal: Leadership in the Three Block War" (Krulak, Marines Magazine, January 1999) succinctly makes this point [5].

The second critical assumption of combat M&S was the concept of aggregation; i.e., that the outcomes of a tactical-level model could be used to inform, or "feed" and operational-level model, and the same concept in turn to use operational-level outcomes to inform theater-level modeling. Thus, the concept of aggregation (with appropriate adjustments for scale and attrition methodologies) enabled the creation of tactical, operational, and theater/strategic models that each provided an adequate representation of the physics-based phenomena associated with the respective level of combat.

It is not clear that this key concept of aggregation will apply to the social science modeling necessary for IW models. Modeling groups of people cannot be done by simply aggregating individual behaviors as if each individual will contribute equally to the behavior of the overall group. While it is clear that the group behavior will be influenced by its members, accounting for the individual contributions is not straightforward. Theories appropriate at a national level, such as macroeconomics, may not be compatible with theories appropriate for a region or town, such as microeconomics. While there are tactical-level actions that impact strategic events in IW that must be represented, in general strategic events are not influenced by the simple aggregation of the tactical and operational actions, as was assumed in DoD combat modeling.

2.1.3 Different Modeling Techniques

Early IW modeling efforts reveal several different modeling techniques that have emerged. Computational social scientists and other modelers believe that agent-based models may prove to be well-suited for representing human behavior. Systems dynamics models, using differential or difference equations, provide a more structured method to examine a complex environment.

Adding human behavior representations to existing combat models is yet another technique used by some. New modeling techniques will bring new challenges to the VV&A processes developed for IW models.

2.1.4 Data

The data required for Irregular Warfare representation is vastly different than the data required for combat models. Performance data (interactions between weapon systems) and scenario data (interaction of weapon systems with the environment) were the two key sets of data needed for combat models. Data that represents how humans interact with each other, how groups interact with each other, and how human perception and behavior change over time due to outside stimuli are just the beginning for IW modeling. Human behavior is influenced by societal and cultural norms, which vary from group to group (nations to neighborhoods). Believing "the Iraqi people" will act as a single entity is as naïve as assuming that Harlem and Manhattan share the same societal and cultural norms because they are both communities in the greater New York City region. The danger in developing new models is that often models are developed with the supposition that the data can be found, formatted, and delivered for any model conceptualized by well-meaning modelers. Many combat models have been shelved because their data requirements could not be fulfilled; either the data did not exist, or obtaining it was cost-prohibitive. As IW models are proposed, data requirements and sources should be demanded of the IW model developers. DoD cannot afford to develop models with an insatiable data appetite.

IW models may be scenario-specific or scenario-independent. Scenario-specific models are developed to address a particular situation; i.e., place, time, actors/subjects. For example, the model SHADOC was developed to examine irrigation management in a specific socio-geographical setting [6]. Scenario-independent models are developed for application to multiple settings. The model provides a particular set of representations and logic, but not tied to any specific place and time, with the goal of being able to address different settings through a change in the data instead of a change to the algorithms encoded in the model. However, since there is such diversity in the fundamental nature of social and cultural factors and establishments in human societies, it is not clear that generalized, data-driven models can be developed for the IW domain.

2.1.5 Uncertainty

IW modeling will need to deal with various sources of uncertainty. First, social science phenomena are not as easily modeled as physics-based phenomena. There are

laws of physics that apply around the globe, and have widely accepted theories, models, and data. Cultural, religious, and social norms are different throughout the world, and these differences have not been exhaustively studied and documented. Thus, the lack of knowledge and data invariably leads to assumptions to fill gaps in understanding that may not be accurate, introducing error into the model. Additionally, a single social science phenomenon may have multiple, conflicting theories that purport to describe it. Reconciling multiple theories may entail developing a capability that represents each theory if sufficient resources permit. Otherwise, accommodations should be made in the code to permit the insertion of a competing theory. The existence of multiple theories underscores an additional source of uncertainty that social science modeling will have.

2.1.6 Fidelity of Results

Because there will be more sources of uncertainty when modeling human behavior, IW models cannot be expected to produce results that have the same level of fidelity as combat models. It is more reasonable to expect IW models to produce a range of possible outcomes or trajectories vice numerical answers with confidence intervals or goodness-of-fit statistics: “Operationally, the most that can be expected is to identify meaningfully different alternative futures and indicators that those alternatives are becoming more or less likely over time” [7]. A reasonable goal for IW models involving human behavior is to be able to identify possible trends, emergent behaviors, and explanatory new theories.

The key differences just discussed serve to illuminate the new challenges that will be presented to DoD as IW models are designed, developed, and fielded. As with any new challenge, opportunity also exists. A disciplined approach to procuring IW models that requires validation to be included in the design and development of new models could prove to be tremendously beneficial in a resource-constrained environment.

2.2 VALIDATION OF IW MODELS

The DoD has spent considerable resources developing the means to verify, validate, and accredit (VV&A) models and simulations used throughout DoD for various purposes. DoDI 5000.61, “DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation” (December 2009) addresses VV&A requirements for M&S used by DoD components [8]. Focusing on validation early in IW model development has the potential to ensure scarce M&S resources are more wisely spent and that the resulting M&S are both relevant and credible. This will be addressed in more detail later in this document.

2.2.1 Social Sciences/Human Behavior

The challenge of validating models of social sciences and human behavior is well recognized. Sokolowski and Banks describe model validation as “the process of comparing simulation results derived from a model against the real-world system that the model is meant to represent” [10]. Further, they state:

The validation of models of physical phenomena is generally straightforward since the laws that govern those systems are usually well known and mathematically precise. In this case, comparing the simulation results against the real-world system is just a matter of matching them to a 100 percent predictable outcome. Validating models of global events containing social components is more problematic [10].

The validation of social science models cannot be done without understanding the model’s conceptual model (a description of what a model (or simulation) must represent; generally thought of in terms of requirement specification) and its referent (a codified body of knowledge about a thing being simulated). For the vast majority of combat models, the conceptual model took the form of a requirements document, and the referent was not specified, because it was implied that the body of knowledge that applied to the requirements was the laws of physics. For social science models needed to represent IW, there are no existing laws to serve as an implied referent. Instead, social science theories provide the referent entailing the body of knowledge that describes particular entity interactions to be represented in the conceptual model. The possible existence of multiple theories that describe a single phenomenon may make the specification of a referent challenging, but specifying a conceptual model and its referent(s) are a necessary and critical step toward the goal of validation of IW models.

2.2.2 Non-aggregation/Complex Behaviors

Irregular warfare typically exists in a nation when there is a strong, armed opposition to the government in power. This points to a disconnect where, at a tactical level, the civilian population in disputed regions of the country are influenced more by government opposition forces than by the ruling government. This phenomenon should force IW modelers to examine a key assumption that was made with combat models: Theater level warfare could be adequately represented by models that aggregated tactical combat engagements into operational campaigns that could, in turn, be further aggregated to represent a theater-wide effort. In practice, this does not hold true for Irregular Warfare. In fact, “What makes IW ‘irregular’ is

the focus of its operations—a relevant population—and its strategic purpose—to gain or maintain control or influence over, and the support of, that relevant population through political, psychological, and economic methods.” [1] Thus, the relationship from tactical to strategic in IW models should be dynamic in the amounts and types of interaction from tactical to strategic as an IW campaign progresses over time. The key validation takeaway is that any social science phenomena that is represented by an aggregation of other social science phenomena should be scrutinized to ensure there is appropriate social science theory that underpins the conceptual model.

2.2.3 Different Modeling Techniques

Agent Based Models (ABM) and Systems Dynamic (SD) models are but two modeling techniques that have come to prominence as DoD has begun to address IW modeling challenges. Both techniques are vastly different from each other, and both are different still from the techniques used to represent the majority of combat models. Often, either because of proprietary reasons, or ignorance, the description of the actual mathematics underlying the modeling techniques cannot be found. A model cannot be validated if it is not transparent and if cause and effect cannot be discerned.

2.2.4 Data

The concept of data validation is not new. Performance data for combat models often was “certified” as appropriate for the model given the purpose of the study. The requirements for IW data are currently unknowable as new models and techniques are still in various stages of development. However, assuming that a single (or several) government organization will be the collector and provider for all IW data needs is naive. Determining the data that will be required for IW modeling, and identifying valid sources for such data, will be a necessary first step in IW model development. But we can assume that some, perhaps the majority, of the data will come from the public domain, and the data will have a relatively short shelf life. (Demographic changes recorded in Baghdad from 2004 to 2009 demonstrate why this is so.) We do not have access to the vast amount of data collected in Iraq, and data gathering efforts in Afghanistan are growing at this time. The physical presence of forces in a country where we are currently conducting operations provides critical data for examining those operations. When examining future operations in a country where we have little presence, such as on the continent of Africa, identifying and collecting the requisite data will be much more challenging. Data is critical to the IW modeling endeavor, and we do not yet know the magnitude and complexity of

this challenge, but demanding that IW model developers identify data requirements, sources and development methods, and then determining the source’s appropriateness for providing useful data will be a necessary first step toward IW data validation.

2.2.5 Uncertainty and Fidelity of Results

One of the keys to validation is ensuring the model outputs reflect reality. Because uncertainty will be greater for the reasons cited in the previous section, it is reasonable to expect that IW model outputs will have greater variance. A key consideration for the validation of IW models should be the manner in which the model is used. Deterministic IW models may be of very little use if they only reflect one possible outcome of millions instead of revealing, of several hundred, or thousand, stochastic model runs, the most likely range of outputs via a response surface methodology. How an IW model accounts for uncertainty, and qualifying the veracity of IW model outputs, should be key elements of the validation of IW models.

3. Best Practices for Validating IW Models

This section presents a framework for the planning and conduct of validation efforts and describes best practices for validation of tactical IW models in the context of that framework. The approach builds upon and extends accepted validation practices to focus on the aspects of IW modeling that distinguish it from traditional combat modeling; in particular, the challenges presented in the previous section. Principal areas of the framework include:

- Identification of user needs.
- Identification of requirements.
- Identification of acceptability criteria traced to the requirements.
- Specification of the intended use.
- Identification of the simuland.
- Identification of referent(s).
- Development of a conceptual model.
- Specification of data (both for assessing and demonstrating model capabilities and for employing the model for a specific purpose).
- Development of the executable model, and generation of model results.

3.1 Validation Framework

To provide a context for description of IW model validation activities, we have extended a diagram developed by Petty from a work entitled “Validation and Verification” by Mikel D. Petty, 2009 [11]. The original is included as Figure 2. The validation framework adds

concepts from DoD best practices and standards [8], [10], to provide the framework shown in Figure 3 in the form of a concept map.

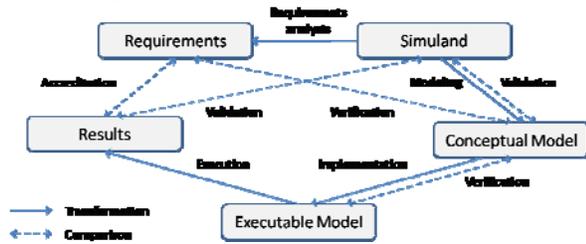


Figure 2. Comparisons in verification, validation, and accreditation [11]

Concepts described in the literature on M&S validation are shown in the boxes; relationships between the concepts are shown as directed arcs to create assertions (e.g., “User Needs specified in Requirements”; “Data defines context for Executable Model”). The framework itself is generic, simply providing a convenient context for discussion of development and validation best practices that can address the challenges presented in the previous section. It would be a serious mistake to ignore the best practices in DoD that have been well established. Instead, we see those practices, including the standardized VV&A documentation specified in MIL-STD-3022 [10], as fundamental to sound validation efforts. Our contribution is in relating these established practices to the additional challenges and requirements of IW model validation. In the following subsections, we describe each of the concepts in the framework and the relationships between concepts, and we identify and discuss best practices for validation of IW models relating to those concepts and relationships.

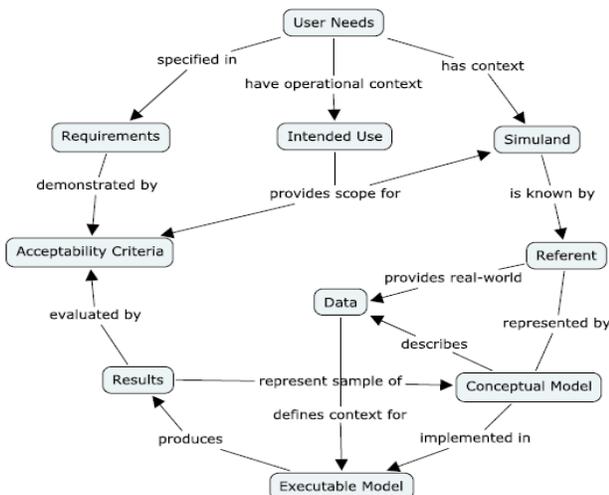


Figure 3. Overview of Validation Framework for IW Modeling Validation Best Practices

3.2 User Needs

Development of any M&S capability begins with a specified user need. The “need statement” initially may not call for an M&S solution explicitly, but may identify an operational capability that is needed to fill a capability gap. In evaluating the need statement, it may be determined that M&S can supply part or all of the required capability. For example, current operations in Iraq and Afghanistan have identified the need for information on population attitudes toward government and coalition operations in order to determine if progress is being made toward achievement of stability and durable peace. These requirements have led to development of the Cultural Geography model by TRAC-Monterey as a partial solution to that need.

Best Practice(s): The developer needs to obtain a succinct and clear statement of the problem the M&S is expected to address. If this information is derived from analysis of an operational need statement, the developer should meet with user representatives to obtain agreement from the users that the proposed M&S capability can address the need, whether through discussions with the user representatives or through more formal contractual negotiations (e.g., proposal evaluation). For IW models, the user need statement must clearly identify the social elements required in the solution (i.e., what aspects of the real world are relevant to the operational situation and user need), any required input sources and fidelity of information required for input to the model, and what information the model needs to produce and at what fidelity.

3.3 Requirements

Requirements specify capabilities and qualities that must be achieved in a solution to meet the user needs. Although requirements analysis and definition are long-established practices in system development, it remains a challenging area. Often, users cannot clearly articulate requirements. Rapid prototyping and other techniques help to provide early implementations to users to help determine if an approach truly addresses the need. The challenges of requirements definition are exacerbated by the complexity and rapid evolution of IW. Best practices in software development (e.g., Carnegie-Mellon University Software Engineering Institute Capability Maturity Model) show that requirements must be well-specified but with recognition that strong change management practices must be in place to manage inevitable change efficiently and effectively.

Concept Map Assertion: *User Needs specified in Requirements.* The requirements analysis process starts with user need statements and creates succinct, specific

statements that can be directly assessed as the development process proceeds.

Best Practice(s): Develop (user and developer working in concert) specific functional or quality statements that can be directly and explicitly assessed to determine if the requirement has been achieved or not in the development of the solution. For IW modeling, this entails being specific about the human or group behavior that needs to be represented. For example, it is not sufficient to say: “Model population attitudes,” but you must specify what portion of the population and what attitudes, with respect to what actions or events, over what period of time, etc.

3.4 Intended Use

The intended use sets the context and scope for a particular model or application of a model. For a particular use, only a portion of the specified requirements may be needed. As stated in MIL-STD-3022 [10], the problem statement serves as the foundation for the definition of requirements, acceptability criteria, and ultimately the accreditation assessment. It documents (1) the question(s) to be answered and the particular aspects of the problem that the M&S will be used to help address; (2) the decisions that will be made based on the M&S results; and (3) the consequences resulting from erroneous M&S outputs. The intended use precisely specifies what the M&S needs to do to support this particular aspect of the user’s need.

Concept Map Assertion: User Needs have operational context Intended Use. Often, models are developed for broad usage. For example, a model such as the Joint Conflict and Tactical Simulation (JCATS) can represent land, air, and sea operations to support joint training. A particular use of the model, however, might be to support training of facility security personnel involving only land-based forces, weapons, and surveillance equipment. That intended use specifies an “operational context” for the model that, in this example, reduces the overall set of capabilities that are needed to satisfy that need. Validation of the model for that intended use would focus on the reduced set of capabilities rather than the entire set of capabilities provided by the model.

Best Practice(s): Obtain a clear, succinct statement of intended use from the user representatives to enable determination of what aspects of the M&S address that particular operational need. For IW modeling, this can entail obtaining or determining details about the scenario to be represented (e.g., population groups, organizations, infrastructure), what functionality will be exercised (e.g., behaviors, interactions) what level of detail is required, and other aspects relating to the social elements of the operations.

3.5 Acceptability Criteria

Acceptability criteria state, for each requirement, conditions by which one can determine if the requirement has been achieved in the developed model, in the context of the intended use (i.e., “a set of standards that a particular model, simulation, or federation will meet to be accredited for a specific purpose”[10]). As discussed above, the particular use to which the model is applied determines the set of requirements that have to be met as well as the criteria against which achievement is evaluated. For example, acceptability criteria establish the level of detail (resolution) and fidelity necessary for the intended purpose, providing a basis for comparison of model execution results to the simuland (that portion of reality to be represented for the intended purpose). Acceptability criteria establish how good is good enough by defining measures of merit, measures of effectiveness, or measures of performance associated for each requirement within the context of the intended use.

Concept Map Assertion: Requirements demonstrated by Acceptability Criteria. Achievement of a requirement for an intended use is demonstrated by satisfaction of the acceptability criteria defined for that requirement and intended use.

Concept Map Assertion: Intended Use provides scope for Acceptability Criteria. The acceptability criteria are established based on a specific intended use for the model. The intended use provides context to scope the acceptability criteria. The intended use scopes the set of requirements and conditions that must be met (e.g., levels of fidelity and detail needed).

Concept Map Assertion: Results evaluated by Acceptability Criteria. The results obtained from execution of the model are evaluated against the acceptability criteria to determine if the criteria have been satisfied for the intended use of the model.

Best Practice(s): Develop a requirements traceability matrix relating each specified requirement with acceptability criteria applicable to the intended use. For each acceptability criterion, define the associated metric that will provide an objective measure of achievement or failure of that criterion. For IW modeling, this can entail describing how modeled behaviors of a population need to compare to social science theory or empirical data (e.g., same causal relations or correlations, acceptable degree of fit to a distribution, etc.).

3.6 Simuland

The simuland is the real-world system of interest, including the objects, processes, or phenomena to be simulated [11]. As such, it is the real-world context for the user needs; i.e., the context within which the user's problem or capability gap exists that requires a solution. For example, the simuland may be a particular type of military operation or mission under a particular set of circumstances and in a particular environment and location. In Figure 2, a principal validation comparison is between the simuland and conceptual model; or, more specifically, between the referent(s) (describing what is known about the simuland) and the conceptual model, as shown in Figure 4.

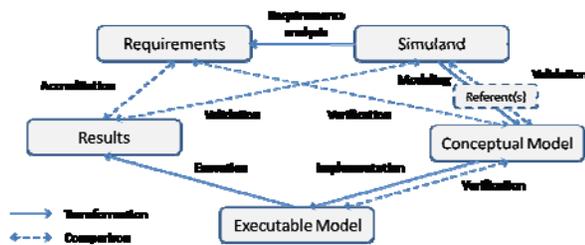


Figure 4. Comparisons in VV&A, with identification of referents (adapted from Petty, 2009) [11].

From our perspective, the simuland relates directly to the user needs, but only to the extent needed based on the intended use.

Concept Map Assertion: User Needs have context Simuland. The simuland is the real-world context for the statement of the operational problem to be solved or operational capability to be provided.

Concept Map Assertion: Intended Use provides scope for Simuland. The NRC report [12] states: "...without a prior specification of intended purpose [use], there is no clear-cut a priori criteria for deciding which features of a phenomenon to stress in its modeled representation." The intended use determines what aspects of the real-world are important to the problem and what can be left out. Statement of intended use is critical for maintaining focus in the application of the model.

Best Practice(s): Derive the simuland from the user needs and intended use to obtain the operational (action-oriented) real-world context for the model. For IW modeling, identify the real-world objects, actors, social relationships, social dynamics, and other aspects of the social dimensions that describe the real-world context for the user needs, but scoped to the intended use.

3.7 Referent

The referent describes what is known about the simuland. For some real-world social science phenomenon, the referent represents a social science theory or collection of empirical data describing that phenomenon. MIL-STD-3022 [10] describes this as a "basis of comparison":

"The basis for comparison serves as the reference against which the accuracy of the M&S representations is measured. The basis of comparison can come in many forms, such as the results of experiments, theory developed from experiments, validated results from other M&S, and expert knowledge obtained through research or from SMEs."

Concept Map Assertion: Simuland is known by Referent. The referent provides the body of knowledge for that portion of the real-world that is of interest for the user needs and intended use (i.e., the simuland).

Best Practice(s): For the various aspects of the real-world objects and phenomena of interest to address the user need in the scope of the intended use, identify the social science theory (or theories, if multiple competing theories will be represented in the model for comparison) that explains that phenomena. Specification of the referent must cite references that describe that theory and its basis for acceptance or examination in the scientific community. Selection of one theory from many available (possibly competing) theories should describe why the particular theory was chosen, what advantages it has (e.g., in availability of supporting data), and other relevant considerations. In the absence of established theory, empirical data may be available that reflects the phenomena of interest. In this case, specify the pedigree of the data (e.g., source, history of use, conditions under which it was collected, etc.). The degree of credibility of the referent within the social science community is an important basis for establishing the validity of the IW model. However, it is also important to note that many areas of social science have competing theories that drive various research efforts. It will generally not be possible to identify a referent in the social science discipline that is universally accepted. This makes specification of the referent all the more important so that the theoretical basis for the model is well-documented and transparent for inquiry. It is also important to note that models can be used to explore new theories. In such a case, the referent may not be an established theory, but an hypothesis regarding the social phenomena of interest, and the purpose of the model is to examine the hypothesis to gain greater insight into the nature of the social phenomena. In this case, the hypothesis and the rationale for advancing that hypothesis may form the referent for the model. In all cases, the best practice calls for explicit specification of the referent for full disclosure of the aspects of the real-

world phenomena that need to be modeled to address the user needs within the scope of the intended use.

3.8 Data

Data is a “representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means.” Data for a model constitutes:

- Parametric data – information that is generally invariant in a model execution, such as (in many cases) equipment performance characteristics or identification of population groups
- Variable data – information that has an initial value but can vary during the model execution, such as level of support to an insurgency group by a particular segment of a population or resources available to conduct a particular operation
- Interchange data – information that is passed to or from the model during execution, either one time or as a continuous stream
- Output data – information produced by the model (referred to as “Results” in Figure 3 and the discussion later in this chapter)

Earlier, we distinguished scenario-specific models from scenario-independent models. Data in scenario-specific models are particular to the setting being represented in the model. The data must therefore be valid not only for purposes of the algorithmic and logical expressions but must also be valid with respect to the specific setting (i.e., correctly represent the features of the setting that are important to the intended use). Scenario-independent models have multiple data sets that must be valid. At least one data set must be appropriate for the algorithmic and logical expressions in order to demonstrate the model meets the requirements for the intended use. Such data does not necessarily have to be an accurate depiction of any “real” setting, but merely valid for use by the algorithms to produce model outcomes that are sufficient for the intended purpose. In contrast, data specific to a particular scenario must be validated with respect to that scenario. For example, if the model will be executed to examine effects of tactical operations on population attitudes in Baghdad circa 2009, then the data must accurately reflect the variables and conditions of interest representing Baghdad in 2009.

Concept Map Assertion: **Referent provides real-world Data.** The referent will identify information that describes the simuland (real-world phenomena of interest). This will include any quantitative and

qualitative data that has been accumulated and is available for use by the model. The referent identifies the data pedigree; i.e., metadata information on the source of the data, processing that has been performed to provide the data, history of use of the data, conditions under which it was collected, etc.

Concept Map Assertion: **Conceptual Model describes Data.** The conceptual model describes the concepts, relationships, and dynamics in some formal way to represent what is known about the simuland in a way that can support the model implementation effort (i.e., software design and development). Fundamentally, the conceptual model describes data and logic. One software engineering principle is to represent as data as much of the model as possible to facilitate later modification (i.e., later changes can be made to the data rather than through software revision, which is generally more costly in time and money). The conceptual model describes the data using various techniques, from very abstract descriptions to highly formalized specifications. Conceptual modeling is discussed in more detail in the next subsection in this chapter.

Best Practice(s): Information provided in the referent needs to be carried over to the data representation in the conceptual model (and, later, the implemented model) to maintain transparency in data source and specification. The data *model* (i.e., specification, schema, design) needs to include metadata to identify the source of the data, how it was collected or generated, its accuracy and precision, type, and legal range of values. The data *content* (i.e., populated data) needs to provide the specified metadata with the data values to support assessment of data quality and credibility along with the use of the information.

3.9 Conceptual Model

The conceptual model is a formal representation of the concepts, relationships, and dynamics identified in the referent(s) to provide a bridge from identification of the real-world phenomena of interest to design and implementation of the executable model. The conceptual model describes the concepts, relationships, and dynamics in some formal way to represent what is known about the simuland to support the model implementation effort (i.e., software design and development). The conceptual model may be expressed in many forms; for example, the Unified Modeling Language (UML), Base Object Model (BOM), or other schemes. Fundamentally, the conceptual model describes data and logic.

As discussed in the Data section above, the conceptual model provides a description of the data input to and used in the model and the data produced by the model to address the intended use. The conceptual model describes

the data vocabulary, structure, and interrelationships, as well as data types, formats, precision, and range of values. The conceptual model stops short of describing a physical data model specific to a particular implementation environment.

The conceptual model also specifies the critical logic to be developed, providing mathematical formulae and other computational descriptions as appropriate to express unambiguously the logic inherent in the referent(s). All constraints relating to the model development are identified; for example, required level of detail in the representation of objects in the model, required performance constraints, and other considerations identified in the requirements. Again, the conceptual model need not express a particular software design, but needs to provide sufficient detail to enable that design to be developed without guesswork regarding what aspects of the simuland are important to the problem. This should be clearly specified in the conceptual model on the basis of the referent(s).

A primary validation activity identified by Petty (2009) [11] is comparison of the simuland and the conceptual model to determine if the conceptual model captures all the key concepts, objects, relationships, and dynamics from the real world necessary to address the intended use for the model. Because referents are descriptions of what is known about the simuland in the context of the intended use, the comparison of interest is more precisely between the referent(s) and the conceptual model.

Concept Map Assertion: **Referent represented by Conceptual Model.** The validation activity confirms that the representation is accurate and complete for the intended purpose.

Best Practice(s): Develop (model developer in close concert with the users) the conceptual model using tools and techniques that create machine-readable specifications of the data and logic of the model (e.g., UML specifications). The conceptual model should provide a platform-independent description of data and logic, leaving implementation issues to the design and development of the executable model (unless such concerns are in some way necessary based on the user need and intended use). Each component of the conceptual model must be traceable to the specified referent(s) to ensure all needed aspects of the social phenomena of interest are represented in the data and logic of the model. It is likely that a conceptual model of complex social phenomena will involve multiple social theories. The combination of these theories into a single model may raise a new concern: how to know if the interplay of multiple social theories is itself a valid representation of the simuland. In such a case, it may be

necessary to bring together expertise to determine if such interaction effects should be added to the referent set for the model as newly proposed or established theory. The key is ensuring transparency and clarity in the conceptual basis for the model.

3.10 Executable Model

Software design and development processes transform the conceptual model into the executable model. The design and development can involve aspects that are not specifically stated in the conceptual model, such as the graphical user interface or execution performance requirements. These may not fall within the purview of the validation effort, in the sense that such functionality or quality characteristics may not be strictly required for the intended use of the model. Maintenance of the requirements traceability matrix throughout the development process will keep such distinctions clearly specified. A key feature of the executable model is its ability to provide trace information to explain cause and effect. Behaviors in IW models often reflect complex combinations of factors and interactions that make it difficult to determine why certain actions occurred, and yet that understanding may be crucial to use of the model to assist decision-makers. In practice, making model execution that transparent is difficult short of tracing execution of the software at the instruction level, which is impractical for model users. The development process should also produce an experimental frame to enable analysts to control the simulation execution, model inputs, and collection and analysis of model results.

Concept Map Assertion: **Data defines context for Executable Model.** The executable model is generally an abstract computational framework permitting instantiation with specific input data. The input data for the execution of the model provides both parametric information characterizing the objects and interactions represented in the model as well as description of the scenario (setting) depicting aspects of the synthetic world that are important to create the operational context for execution of the model. Note that in some cases the parametric information may apply across scenarios; that is, the data may be scenario-independent, such as the performance characteristics of a vehicle or weapon system or even the demographics of the population in a region of interest. In other cases, the parametric data may be scenario-specific, having certain value settings because of the conditions being represented in the particular scenario.

Concept Map Assertion: **Conceptual Model implemented in Executable Model.** The software design and implementation processes transform the conceptual model into an executable model. Verification is the

principal evaluation methodology to assess the correctness of that transformation.

Best Practice(s): Design the model implementation to be as transparent as possible to permit analysis of execution paths and computed outcomes. The executable model should aid analysts in tracing cause and effect (“Why did this outcome occur?”) or determining correlations of results to input factors. Design and develop an experimental frame to facilitate set-up and conduct of validation tests supporting examination of the computational characteristics of the model.

3.11 Results

The results represent information generated by the model. Such information may entail data presented to the user interface, data stored to persistent storage, and data sent to other systems through some data interchange mechanism. The results represent a particular sample from the set of all possible outputs that can be generated by the executable model through its various logical and computational processes (whether stochastic or deterministic) operating on the full range of data inputs. From (Petty, 2009) [11], a key validation effort is comparison of the simulation results to the simuland; however, our framework takes the approach that the results are evaluated against the acceptability criteria. The acceptability criteria specify explicitly how to determine if the results are “good enough” for the intended use. If the requirements have been carried through to identification of the simuland, referent(s), and conceptual model, under the scoping of the intended use, then comparison of the results to the acceptability criteria provides a direct assessment of the “goodness” of the model for the intended use and “closes the loop” as illustrated in Figure 3.

Concept Map Assertion: **Executable Model produces Results.** Clearly, the results are the product of the computations performed by the executable model.

Concept Map Assertion: **Results represent sample of Conceptual Model.** The execution of a model produces one outcome, one sample, of the set of possibilities represented in the logical specification given in the conceptual model and implemented in the executable model.

Best Practice(s): The acceptability criteria identify what the model needs to do to satisfy or meet the set of respective requirements pertinent to the intended use. The acceptability specify the conditions to be met and the quantitative and qualitative metrics used to measure their success [10]. The validation tests identified for the acceptability criteria identify all expected results for each

test case. The principal validation practice, therefore, is comparison of actual computed results with the expected results, with analysis of discrepancies to determine if the differences are due to a failure in the implementation, a failure in the conceptual model, a failure in specification of the test, or a failure in specification of the requirement itself. If expected results are described in terms of distributions, then the model needs to be run over the set of inputs to generate multiple samples to enable analytical comparison of the actual distribution of outcomes to the expected distribution of outcomes. The fundamental theme through this comparison and analysis, as in the other elements of the validation framework, is transparency in describing the model results compared to the expected results.

4. Summary and Way Forward

The purpose of the IW MMT validation best practices guide (BPG) is to provide a clear framework for the same. The aim of this paper is to share some of the insights from the BPG, namely the sections that define IW and the modeling challenges that IW analysis presents, as well as the section that enumerates the validation framework in detail. Additional efforts for the TRAC-MTRY/NPS validation effort include application of this best practices guide to IW MMT currently used in the modeling and simulation community. The MMT to be utilized for this application are the Cultural Geography (CG) Model, developed by TRAC-MTRY and NPS, and the Peace Support Operations Model. At the time of completion of this paper, the application of the BPG to the selected MMT is in progress.

4. References

- [1] Department of Defense. *Irregular Warfare Joint Operating Concept, v1.0*. September 11, 2007.
- [2] Department of Defense. *Directive 3000.07 Irregular Warfare*. December 1, 2008.
- [3] Department of the Army. *FM 3-24/MCWP 3-33.5, Counterinsurgency*. 2006.
- [4] Epstein, Joshua M., and Axtell, Robert. *Growing Artificial Societies: Social Sciences from the Bottom Up*. Brookings Institution Press. 1996.
- [5] Krulak, Charles C. *The Strategic Corporal: Leadership in the Three Block War*. Marines Magazine. January 1999.
- [6] Barreteau, O. and Bousquet, F. *SHADOC: a multi-agent model to tackle viability of irrigated systems*. *Ann. Operation Research* 94, pp. 139–162. 2000.
- [7] Alberts, David S., and Hayes, Richard E. *Understanding Command and Control*. DoD Command and Control Research Program. CCRP Publications. Washington, D.C. 2006.
- [8] Department of Defense. *DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)*. Department of Defense Instruction 5000.61. USD(AT&L). December 9, 2009.
- [9] Sokolowski, John A., and Banks, Catherine M. *Modeling and Simulation for Analyzing Global Events*. John Wiley & Sons, Inc. Hoboken, NJ. 2009.
- [10] Department of Defense. *Department of Defense Standard Practice: Documentation of Verification, Validation, and Accreditation (VV&A) for Models and Simulations*. MIL-STD-3022. January 28, 2008.
- [11] Petty, Mikel D. (2009). “Verification and Validation.” *In Principles of Modeling and Simulation: A Multidisciplinary Approach*. Edited by John A. Sokolowski and Catherine M. Banks. John Wiley & Sons, Inc. Hoboken, NJ.
- [12] NRC (National Research Council). *Behavior Modeling and Simulation – From Individuals to Societies*. G. L. Zacharias, J. MacMillan, and S. B. Van Hemel (Eds.). Committee on Organization Modeling: From Individuals to Societies. National Academies Press. Washington, DC. 2008.

Acknowledgements

Opinions expressed in this paper are those of the authors and do not necessarily represent the position of any of their respective organizations or SISO.

Authors' Biographies

CPT RICHARD F. BROWN is the TRAC-MTRY lead for the IW validation best practices effort.

CURTIS BLAIS is a Research Associate in the Naval Postgraduate School Modeling, Virtual Environments, and Simulation (MOVES) Institute. Principal research interests include the application of current and emerging Semantic Web technologies to obtain valued information for military decision-makers in net-centric architectures and to improve interoperability across C2 and M&S systems, and advancing the state of M&S to address Irregular Warfare. Mr. Blais has a B.S. and M.S. in Mathematics from the University of Notre Dame.

DR. JEFF APPEGET is a Senior Lecturer in the NPS Operations Research Department. He retired as an Army Colonel with 30 years of service, 20 of those as an Army OR analyst. He leads a research initiative at NPS for the Human, Social, Culture and Behavior (HSCB) Modeling Program, and teaches the Wargaming Analysis course at NPS. His research interests include Irregular Warfare and Counterinsurgency, Wargaming, and Integer Programming. He holds a BS from the US Military Academy (1979), MS in Operations Research and Statistics from Rensselaer Polytechnic Institute (1989) and a PhD in Operations Research from the Naval Postgraduate School (1997).

TIMOTHY PERKINS is a Research Community Planner with the U.S. Army Engineer Research and Development Center (ERDC). Mr. Perkins and conducts interdisciplinary research primarily pertaining to infrastructure, essential services and agent-based modeling and simulation, infrastructure, essential services and populations. Mr. Perkins is currently on an assignment as a Visiting Research Analyst at with the U.S. Army TRADOC Analysis Center (TRAC)-Monterey, located at the Naval Postgraduate School. He received a Master of Urban Planning degree and B.S. in Advertising from the University of Illinois.

DR. DEBORAH DUONG is a Senior OR Analyst at Augustine Consulting, Inc. on contract at US Army TRAC-Monterey. Dr. Duong has nearly twenty years experience in Computational Social Science, including invention of the first agent based cognitive social model in 1991. Her experience includes OSD HBR modeling on the Joint Warfare Simulation (JWARS) project and for the IW team, and US Army INSCOM work on data representation. She authored the Nexus, Oz and Indra programs, and is now working on US Army TRAC's Cultural Geography model.