TRAC-M-TR-15-013
March 2015

Modeling and Simulation Behavior Validation Methodology and Extension to Model Validation for the Individual Soldier

TRADOC Analysis Center
700 Dyer Road
Monterey, CA 93943-0692

This study cost the Department of Defense approximately $179,420 expended by TRAC in Fiscal Year 13 - 15.
Prepared on 20150225
TRAC Project Code # 060033

THIS PAGE INTENTIONALLY LEFT BLANK

TRAC-M-TR-15-013
March 2015

Modeling and Simulation Behavior Validation Methodology and Extension to Model Validation for the Individual Soldier

Authors

MAJ Adam Haupt
Dr. Thomas Anderson
LTC William Platte
LTC Thomas Deveans

PREPARED BY:

ADAM HAUPT
MAJ, US Army
TRAC-MTRY

APPROVED BY:

CHRISTOPHER M. SMITH
LTC, US Army
Director, TRAC-MTRY

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** 25 March 2015	**3. REPORT TYPE AND DATES COVERED** Technical Report, June 2013 to December 2014

4. TITLE AND SUBTITLE
Modeling and Simulation Behavior Validation Methodology and Extension to Model Validation for the Individual Soldier

5. PROJECT NUMBERS
TRAC Project Code 060033

6. AUTHOR(S)
MAJ Haupt, Dr. Anderson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army TRADOC Analysis Center - Monterey
700 Dyer Road
Monterey CA, 93943-0692

8. PERFORMING ORGANIZATION REPORT NUMBER
TRAC-M-TR-15-013

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
TRADOC Analysis Center – Methods and Research Office (MRO)
Army Research Labs (ARL)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
TRAC-M-TR-15-013

11. SUPPLEMENTARY NOTES
Findings of this report are not to be construed as an official Department of the Army (DA) position unless so designated by other authorized documents.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)
The modeling and simulation (M&S) knowledge base does not completely describe or document the validation of individual model behaviors for rapid searching, implementation, and distribution to the M&S community. This results in a M&S development environment where the usage of potentially flawed or inaccurate behaviors, vignettes and models is possible. In FY 2014 the Army Research Lab (ARL) and TRAC WSMR sponsored TRAC MTRY to create a methodology for the validation and documentation of M&S behaviors. The resulting methodology was a validation process that would define, describe and characterize behaviors and provide a validation score that addressed the realism of that behavior. This was accomplished with a validation meta-model and a supplementary scoring model that was composed of four criteria; conceptual validation, operational validation, subject matter expert and documentation.

14. SUBJECT TERMS
Modeling and Simulation, Combat Model, Behavior Validation, Operational Validation, Conceptual Validation, COMBATXXI, Meta-Model

15. NUMBER OF PAGES
67

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

NOTICES

DISCLAIMER

Findings of this report are not to be construed as an official Department of the Army (DA) position unless so designated by other authorized documents.

REPRODUCTION

DISTRIBUTION STATEMENT

Approved for public release; distribution is unlimited.

DESTRUCTION NOTICE

When this report is no longer needed, DA organizations will destroy it according to procedures given in AR 380-5, DA Information Security Program. All others will return this report to Director, TRAC, ATTN: ATRC, 255 Sedgwick Avenue, Fort Leavenworth, Kansas 66027-2345.

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The modeling and simulation (M&S) knowledge base does not completely describe or document the validation of individual model behaviors for rapid searching, implementation, and distribution to the M&S community. This results in a M&S development environment where the usage of potentially flawed or inaccurate behaviors, vignettes and models is possible. In FY 2014 the Army Research Lab (ARL) and TRAC WSMR sponsored TRAC MTRY to create a methodology for the validation and documentation of M&S behaviors. The resulting methodology was a validation process that would define, describe and characterize behaviors and provide a validation score that addressed the realism of that behavior. This was accomplished with a validation meta-model and a supplementary scoring model that was composed of four criteria; conceptual validation, operational validation, subject matter expert and documentation.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AR	Army Regulation
ARL	Army Research Labs
BV	Behavior Validation
BVM	Behavior Validation Model
BOCME	Basic Observation, Communicate, Move and Engage
COP	Combat Outpost
ERDC	Engineer Research and Development Center
HBM	Human Behavior Model
HTN	Hierarchal Task Network
M&S	Modeling and Simulation
MTRY	Monterey
SME	Subject Matter Expert
SoS	System of Systems
TRAC	Training and Doctrine Command Analysis Center
TRADOC	Training and Doctrine Command
UAS	Unmanned Aerial System
VRDM	Vector Relational Data Model
VV&A	Verification, Validation and Accreditation
WSMR	White Sands Missile Range

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would especially like to thank Mr Paul Works at the TRAC-MRO for making this project possible with his guidance and resourcing. I would like to recognize Dr. Thomas Anderson, for his tireless efforts above and beyond what was asked of him in support of this effort. Additionally, I would like to thank Mr. Chad Mullis and his team for their invaluable help and guidance in creating a relevant product that will support the important analysis work that they do for America's warfighters.

THIS PAGE INTENTIONALLY LEFT BLANK

SECTION 1. INTRODUCTION

1.1. PURPOSE

The purpose this project is to develop a methodology to validate and document simulation model behaviors in COMBATXXI in order to improve the ability of users to reuse valid behaviors and add to the pedigree of study results that these behaviors contribute to. This metadata model will provide a framework that may provide for improved transparency for the realism of individual model entity behaviors through the creation of an easily searchable knowledge base which can be dynamically expanded to meet the growing simulation needs of the military. The intent is that this framework will be initially applied to COMBATXXI behaviors, but will be robust enough to be applied to any current or future simulation model behavior.

1.2. BACKGROUND

The modeling and simulation (M&S) knowledge base does not completely describe or document the validation of individual model behaviors for rapid searching, implementation, and distribution to the M&S community. This results in a M&S development environment where the usage of potentially flawed or inaccurate behaviors, vignettes and models is possible. In FY 2014 the Army Research Lab (ARL) and TRAC WSMR sponsored TRAC MTRY to create a methodology for the validation and documentation of M&S behaviors.

1.2.1. Simulation Models and Validation Overview

Simulation models are software implemented mathematical or logical representations of real-world systems and processes (Law, 2007). The military increasingly uses simulation models, such as COMBATXXI to solve problems and to aid in decision-making.

> The developers and users of these models, the decision makers using information obtained from the results of these models, and the individuals affected by decisions based on such models are all rightly concerned with whether a model and its results are ‘correct’ (Sargent, 2011, p. 183).

Models are comprised of multiple interactions of entity behaviors, where more complex behaviors are an aggregation of simpler or ‘primitive’ entity behaviors. The resulting analytical insight derived from any model or simulation is only as good as the behaviors within. Model validation is defined as “the process of determining the extent to which the M&S adequately represents the real-world from the perspective of its intended use” (AR 5-11.30, 2005). Currently there is not a universal framework that rigorously defines M&S behavior validation data, quantifies the validity of behaviors, and offers model transparency to the behaviors. However, by addressing the key elements of M&S validation we are able to create a framework to apply to individual model behaviors.

1.2.2. COMBATXXI Overview

COMBATXXI is the primary combat simulation model used at TRAC WSMR to analyze combat operations. COMBATXXI is:

- A joint, high-resolution, closed-form, stochastic, discrete event, entity level structure analytical combat simulation.
- Developed and supported by the TRAC-WSMR and partner organizations.
- Designed for simulation of operations at the brigade level or lower with appropriate representation of joint/higher echelon assets.
- Used for land and amphibious warfare analyses in the Research, Development and Acquisition and Advanced Concepts and Requirements M&S domains.

Major model functions include:

- Ground combat: Light and heavy forces.
- Air mobile forces.
- Future forces.
- Fixed-wing and rotary-wing: Close air support, armed reconnaissance, detailed communications modeling, rotary-wing, and direct/indirect fire.
- Amphibious operations.

- Urban operations.
- Combat service support - logistics and casualty handling (TRAC WSMR, 2014).

1.2.3. Simulation Model Behaviors

COMBATXXI defines behavior as a set of instructions that simulates the operation of scenario elements and the decision making process. A behavior can be formulated by an internal or a user defined process. Element operations can stimulate effects or respond to stimulus from other elements in the operational environment (TRAC WSMR, 2012).

1.2.3. Current Behavior Validation Methods in COMBATXXI

Currently, there is a limited knowledge base on the validity of individual model behaviors in COMBATXXI. The primary library for COMBATXXI behaviors is the COMBATXXI Scenario Library, which catalogues critical scenario components, to include model behaviors. The COMBATXXI Scenario Library is updated and disseminated with each new updated release of COMBATXXI. Behaviors are rated on a three star scale. Stars are awarded to behaviors to reflect how complete the behavior metadata is filled out and how much additional supporting documentation is present. For example, a one star rating means that the behavior has a Point of Contact, Keywords and a Description. Whereas a two star behavior would have the same fields filled out as a one star behavior, but in addition would contain additional documentation that would explain how the behavior conceptually works (e.g., flowcharts) and perhaps a reference to a study that the behavior was used in (TRAC WSMR, 2014). This rating method is very useful for users who are looking for well documented behaviors so that they can reuse a given behavior in a new study, but it does not specifically speak to how realistic a model behavior is in comparison to the real life system.

SECTION 2. METHODOLOGY

2.1. CONSTRAINTS, LIMITATIONS, & ASSUMPTIONS

TRAC defines constraints as restrictions imposed by the sponsor that limits the research team's options in conducting the research. TRAC defines limitations as an inability of the research team to fully meet the research objectives or fully investigate the research issues. TRAC defines assumptions as statements related to the research that are taken as true in the absence of facts, often to accommodate a limitation. The following constraints, limitations, and assumptions were developed when modeling force protection technologies:

Constraints:

- Study must be completed no later than 31 December 2014.
- Entity behavior limited to select COMBATXXI Mobility, Unmanned Aerial System (UAS) and human behaviors.

Limitations:

- Number and type of available scenarios and behaviors are limited.

Assumptions:

- TRAC-MTRY analysts are sufficient Subject Matter Experts (SME) for the purpose of validating Human Behavior in COMBATXXI.
- Current COMBATXXI baseline behavior documentation will be made available.
- Scenarios provide sufficient detail to assess effectiveness of TRAC validation framework.

2.2. APPROACH

The project team followed the approach outlined in Figure 1 to address the problem statement. After defining the problem, the research team conducted a literature search on how COMBATXXI classifies entity behavior, doctrinal definitions of the Army's Warrior Tasks, and how the computer modeling and psychology community define behavior. Next, the research team reviewed current model simulation validation approaches and techniques to find a suitable validation methodology to apply to the problem statement.

A TRAC M&S behavior validation methodology was developed by applying the model validation concepts described above to M&S entity behaviors. This entailed developing a meta-model framework with robust behavior descriptors and information relationships that allow model transparency. A scoring methodology for behavior validity was created to provide a means of characterizing behaviors. The TRAC behavior validation framework and methodology was then implemented for selected COMBATXXI behaviors to test the effectiveness of this construct. Following effective testing the results were documented, the validation methodology was finalized and conclusions were made.

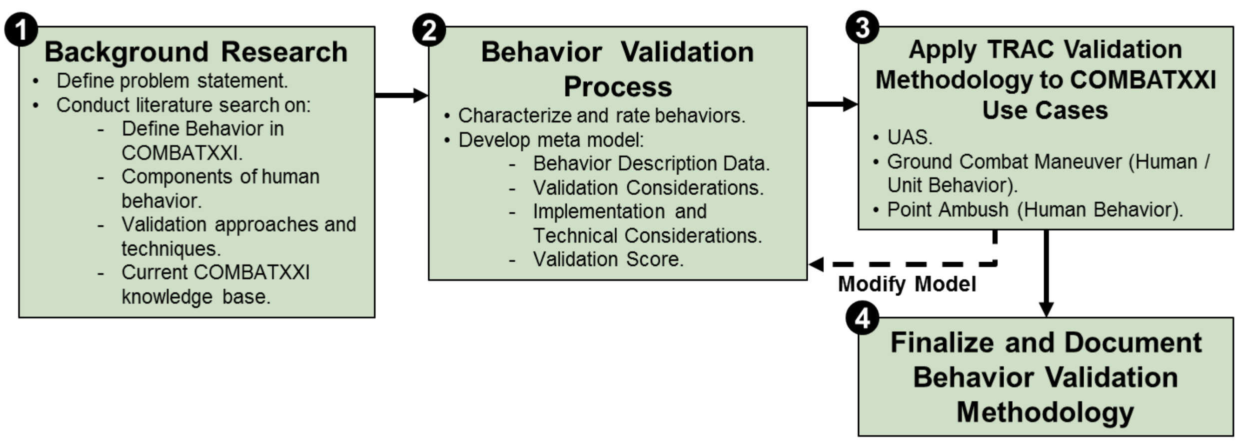


Figure 1: Behavior Validation Methodology.

2.3. STUDY TEAM

- MAJ Adam Haupt, Combat Analyst, TRAC-MTRY.
- Dr. Thomas Anderson, ERDC Analyst, ERDC.
- LTC William Platte, Knowledge Management Advisor, TRAC-MTRY.
- LTC Thomas Deveans, Combat Analyst, TRAC-MTRY.

SECTION 3. VALIDATING MODEL BEHAVIORS

3.1. BEHAVIOR SCORING APPROACH

A scoring model validation approach was used as the basis for developing our behavior validation methodology. In this validation approach a scoring model is used to calculate the validity of the model behavior. Scoring criteria are developed and weights are subjectively applied to those criteria by a validation team (Sargent, 2011).

This project team decided that a scoring model implemented by a validation SME or team would be the best approach for documenting and validating model behaviors. The advantage of a scoring model is that it would quickly separate behaviors by degrees of acceptability inside a knowledge management system allowing developers to prioritize and identify behaviors that need improvement or were sufficient for their current needs. An additional advantage is that users would be able to quickly search for highly rated behaviors when creating new combat scenarios. Finally, the scoring method allowed this project team to evaluate behaviors by criteria that we viewed as relevant and practical to the military.

3.2. META-MODEL DEVELOPMENT

The study team created a behavior validation methodology by designing a meta-model framework that contains all necessary metadata that describes, classifies and ultimately validates each behavior. The meta-model had four broad categories. 1) Behavior Description Data – Generalized data that classifies and describes the behavior. 2) Validation Considerations – Information that was considered in the validation of each behavior. 3) Implementation and Technical Considerations – Information and documentation critical to understanding how the behavior works, how it has been used and how it is implemented. 4) Validation Score – A validation score developed using a scoring model developed by this study team to rate the overall level of validity that a behavior demonstrated.

3.2.1. Behavior Description Data

The Behavior Description Data consists of the critical metadata that describes and classifies the model behavior. There are nine data fields for the behavior description portion of the meta-model:

1. Behavior Name.
2. Date (created/updated).
3. Behavior Description: Brief description of what the behavior does, to include trigger and action.
4. Parent Behavior: List of parent behaviors that use the named behavior as a component.
5. Component Behaviors: List of required component behaviors that are required for the named behavior to work properly.
6. Associated Vignettes: List of vignettes that the named behavior has been used in.
7. Key Entities: Entity types that the named behavior has been applied to.
8. Behavior Resolution/Type: The level that the named behavior is applied to (Functional Capability, Entity, Unit, Operational).
9. Behavior Type: Type of behavior that speaks to behavior complexity (Cognitive, Tactical, Procedural, Primitive).

The most important element of this descriptive data is the classification of the model behavior. We initially identified three behavior resolutions, which were Entity, Unit and Operational from COMBATXXI documentation. After periodic In-Progress Reviews, TRAC WSMR requested that we add a fourth behavior resolution, based on their growing practice of creating behaviors for functional capabilities or subsystems of entities. The resulting four behavior resolutions in COMBATXXI are:

1. Functional Capability (System / Subsystem) – Behavior applied to a specific entity capability such as a sensor or a weapons control system.

2. System/Entity Behavior – Behaviors for single model entities (e.g., Movement, Weapons Engagement, Civilian Behavior).

3. Unit Behavior – Behaviors for multiple entities working as a unit (e.g., Maneuver, Coordinated Engagements).

4. Operational Behavior – Behaviors designed for command and control of multiple units (ie. Coordination of battalion and above assets such as Fires, ISR, Coordinated Maneuver between units).

These behavior resolutions are subject to multiple types of behaviors. These types are Cognitive, Tactical, Procedural and Primitive, illustrated in Figure 2. A cognitive behavior is the all-encompassing mental process associated with problem solving. In a COMBATXXI combat environment entities execute multi-function tactical decisions that are made up of simpler procedures that are doctrinally sound.

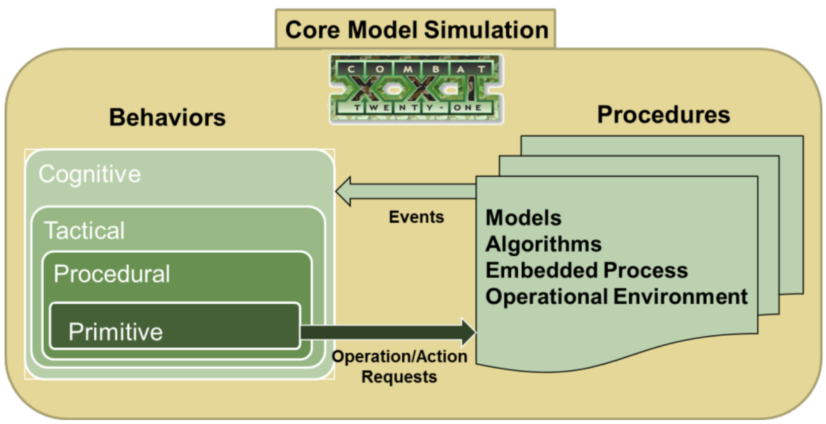


Figure 2. Layered Relationship Between Behavior Types from (TRAC WSMR, 2012).

These procedures are in turn made up of a collection of primitive behaviors that are considered the simplest building blocks of any entity behavior. These layered behaviors result in a decision which is manifested as an action in the COMBATXXI environment. These actions feed additional stimuli to the acting entity which loops the process back to the cognitive decision making process. An example of how behavior layers are tied to functional behavior models is illustrated in Figure 3, which illustrates some of the functions associated with each behavior type and examples of corresponding behavior functions in COMBATXXI (TRAC WSMR, 2012).

Behavior Type	Function	Example
Cognitive Behavior	• Capture & Process Information • Tactical Decision Making. • Operates elements. • User controlled.	BSL Behavior : Engage Targets Trigger Event: Potential Target List Complete 1: BSL Cmd: PermitDirectFireEngagements 2: BSL Cmd: PassPotentialTargetsToEngage
Tactical	Execute Entity Actions to Multi-Function Tactics. • Decides how to process action. • System controlled	1: TargetingDMI2.setIsPermittingEngagement(T) 2: TargetingDMI2.detect()
Procedural	Single Function Tactical Actions. • Routes action to proper FM. • System controlled.	1: EngageDMI.setHoldFire(); 2: EngageDMI.beginEngagement
Primitive	• Low-Level, Basic Entity Function. • Implements FMs • Access physical models. • System Controlled	1: EngageFM.setHoldFireStatus() 2: EngageFM.beginEngagement()

Figure 3. Layered Behaviors Example from (TRAC WSMR, 2012).

3.2.2. Validation Considerations

The Validation Considerations portion of the meta-model contains six data fields that are the basis for feeding the validation scoring model. These fields highlight the 'lens' that the validation team evaluated the behavior through and additionally provide important fields that help users to conduct more specialized searches for behaviors that they may need for future scenario development. These validation data fields are:

1. Validation Technique: Academically accepted techniques for validating simulation models. While there are 15 documented validation techniques that we found in our literature review, see Appendix D. Validation Techniques, we chose to initially select five that seemed to be easily applied to COMBATXXI:

    a. Animation: "Operational behavior is displayed graphically as the model moves through time" (Sargent, 2011) (ADRP 3-0, 2012).

    b. Comparison: Behavior inputs and outputs are compared to preexisting behaviors that are viewed as reasonably valid.

c. Event Validity: Behavior output is compared to the data from the real system.

d. Face Validity: The most common form of validation, which involves a subject matter expert assessing the inputs and/or outputs of the behavior and making a subjective judgment on its validity based off of experience or professional knowledge.

e. Traces: Behavior outputs of specific entities are followed (traced) through the model to determine if the model's logic is correct or accurate.

2. Warfighter Function: COMBATXXI is primarily used as a combat simulation model and consequently most of the behaviors are combat related or are necessary components of a combat behavior. In order to speak to the realism of those combat behaviors we used the Army Warfighter Functions as a validation consideration because it allowed the behavior realism to be compared to its adherence of Army doctrine and combat performance, with the added benefit of allowing more specialized searches by a user looking for a behavior that fit in one of the Army Warfighting Functions. The six Army Warfighting Functions are:

    a. Mission Command: "The mission command warfighting function is the related tasks and systems that develop and integrate those activities enabling a commander to balance the art of command and the science of control in order to integrate the other warfighting functions" (ADRP 3-0, 2012, pp. 3-2).

    b. Movement and Maneuver: "The movement and maneuver warfighting function is the related tasks and systems that move and employ forces to achieve a position of relative advantage over the enemy and other threats. Direct fire and close combat are inherent in maneuver" (ADRP 3-0, 2012, pp. 3-3).

    c. Intelligence: "The intelligence warfighting function is the related tasks and systems that facilitate understanding the enemy, terrain, and civil considerations" (ADRP 3-0, 2012, pp. 3-4).

d. Fires: "The fires warfighting function is the related tasks and systems that provide collective and coordinated use of Army indirect fires, air and missile defense, and joint fires through the targeting process" (ADRP 3-0, 2012, pp. 3-4).

e. Sustainment: "The sustainment warfighting function is the related tasks and systems that provide support and services to ensure freedom of action, extend operational reach, and prolong endurance" (ADRP 3-0, 2012, pp. 3-4).

f. Protection: "The protection warfighting function is the related tasks and systems that preserve the force so the commander can apply maximum combat power to accomplish the mission" (ADRP 3-0, 2012, pp. 3-4).

3. Human Behavior: This is a binary (Yes/No) field that allows the developer to indicate that the named behavior was designed for humans and should be evaluated by how well it performs to the human behavior components listed below.

4. Human Behavior Component: This data field was created to address Army Research Lab's interest in validating human behavior. We realized that not all behaviors are inherently combat related (e.g., Civilian urban foot traffic behavior during a 24 hour cycle). Additionally, there are some combat behaviors that are influenced more by the human condition than by doctrinal training in some circumstances (e.g., Soldier's ability to detect audible movements of enemy troop formations after significant numbing of senses after artillery barrage). There are six human behavior components that we identified from literature review that are accounted for in our behavior validation model (Silverman, Bharathy, O'Brien, & Cornwell, 2006):

    a. Affect: Emotion or attitude generated from stimuli.

    b. Perception: Organization, identification, and interpretation of sensory information in order to represent and understand the environment.

c. Cognition: Mental process related to problem solving and decision making.

d. Biology: Physical and mental capabilities and limitations (e.g., effects of stress and fatigue on Soldier, weight and speed limitations).

e. Social: Effects of social norms and pressures on human behavior.

f. Memory: Process of storing and recalling information.

These are the essential elements for a validation framework of human behavior in modeling and simulation; however, human behavior subject matter experts must determine the extent to which these aspects of human behavior comprise a behavior and are well represented.

5. Reference Doctrine: List of doctrinal reference that the named behavior was designed to support and validated against for operational realism.

6. SME: List of SMEs that were responsible for the development and validation of the named behavior. Level of expertise and experience is used to add weight to the validation score of the named behavior.

3.2.3. Implementation and Technical Considerations

The Implementation and Technical Considerations portion of the meta-model consists eight fields that address the implementation methods used to create the named behavior, the required software versions and requirements, the supporting documentation that help implement and validate the named behavior and the security considerations. This element of the meta-model has some bearing on the validation score because the level of documentation weights into the score, but is primarily used to ensure that users are aware of the technical requirements and security/distribution constraints associated with the named behavior. The eight fields in the portion of the meta-model are as follows:

1. Implementation Method: Description of the coding language or tool that the named behavior is implemented in (e.g., BSL, Python, Hierarchal Task Network (HTN).

2. Project Study: List of studies that the behavior has been used in. Allows user to see examples of how the behavior has been used before and what types of results it helped yield.

3. Reference Model Implementations: List of modeling tools that the behavior was implemented in. For the purposes of this study the modeling tool is COMBATXXI, but this methodology could be expanded to any combat model (e.g., AWARS, OneSAF, etc.).

4. Reference Model Version: Version of the model tool. Important because not all behaviors are necessarily compatible with different versions of COMBATXXI.

5. Reference Documentation: List of reference documentation that would ideally hyperlink to the digital document. Increases the academic pedigree of the behavior. There are three classifications of documentation prescribed in the SITS Library which is published on their CONFLUENCE website (TRAC WSMR, 2014):

    a. Documentation: Material that describes the conceptual model and how to code and implement. Should speak to what degree the conceptual model reflects the real life system.

    b. Confirmation: Material that proves that the behavior works correctly from a technical aspect.

    c. Verification: Material that verifies that the behavior outputs are realistic.

6. POC: List of persons to contact for more information on the named behavior.

7. Security Classification: Level of classification that the behavior was designed for and can be used in.

8. Distribution: Field contains the distribution limitations that the owning organization places on the named behavior.

3.2.4. Validation Score

The end result of our behavior validation methodology is the final validation score which is computed from a scoring model. The scoring model, as earlier mentioned, is based off of one of the common validation approaches used in M&S. It is worth mentioning that there are four basic approaches for validating a model behavior or simulation model (Sargent, 2011).

1. Model development team itself makes a subjective decision as to whether the simulation model is valid. This is based off of results from various tests and evaluations conducted as part of the model development process.
2. Users, who are heavily involved with the model development, determine the validity of the simulation mode. This method is preferred when the modeling team is small, because it adds credibility to model.
3. An independent third party team accesses the simulation model's validity. This approach is generally called "independent verification and validation" (Sargent, 2012). This approach should be used when developing large-scale simulation models, whose developments usually involve several teams.
4. A scoring model is used to calculate the validity of a simulation model. In this approach scoring criteria are developed and weights are subjectively applied to those criteria by a validation team.

We chose the scoring model approach because it could quickly separate behaviors by degrees of acceptability inside a knowledge management system allowing developers to prioritize and identify behaviors that need improvement or were sufficient for their current needs. An additional advantage is that users would be able to quickly search for highly rated behaviors when creating new combat scenarios. Finally, the scoring model approach allowed this project team to evaluate behaviors by criteria that we viewed as relevant and practical to the military.

3.2.4.1. Initial Scoring Model

Our initial scoring model had four components that we decided where relevant for grading the validity of a behavior. These four components are depicted in Figure 4.

Weight	Criteria	Description
0 - 4	Doctrine / Realistic	Behavior supported by Doctrine or is realistic.
0 - 2	Documented	Behavior algorithm was documented in reports/pubs.
0 - 2	SME Developed	Behavior was created by a credible SME.
0 - 2	Reusable	Behavior can be reused in multiple CXXI scenarios.

Validation Scale		
0 - 3	4 - 7	8 - 10

Figure 4. Initial Scoring Model Components and Scale.

In this model we decided that Doctrine/Realistic and SME Developed both spoke to the overall realism of the behavior and that it accounted for potentially 60% of the overall score. The Documented component was meant to capture the degree to which the behavior was bolstered by relevant documents and analysis to increase its academic pedigree and the Reusable component was thought to be a valuable component to allow users to decide if the behavior was reusable enough to warrant implementation into a new scenario. The we divided the final validation score into three levels; red, amber and green, which in common Army usage equates to; not valid, needs improvement and valid. At this early developmental phase this seemed sufficient to categorize behaviors and focus an organization's limited resources on prioritizing further development efforts on the category of behaviors that held the lowest validity. After initially proposing this scoring model to the TRAC Board of Directors, TRAC WSMR and ARL we made changes to this model based off of common critiques. First, all parties recommended greater weight and specificity to the components that evaluated the realism of the behavior. Second, all parties required better definition and specificity of the Documented component and how it would be scored. Third, the Reusable component was deemed unnecessary in validating behaviors, although it was agreed that it was desirable for a behavior to have this quality. Forth, greater definition of the final validation score needed to be addressed. With these critiques we created an updated scoring model to address these concerns.

3.2.4.2. Final Scoring Model

Our final behavior validation scoring model was changed to reflect the criticisms addressed previously and to more accurately adhere to the framework of the modeling process in **Figure** 5. Our scoring model has four components that are weighted and scored to yield a validation score, see . Those components are:

1. Conceptual Validation – Scored on an ordinal scale of 0-2 and then multiplied by 2. The degree of realism and adherence to modern military doctrine that the behavior's conceptual model demonstrates. This directly refers to the conceptual design and inputs of the behavior and how it is designed to work (Sargent, 2011).

2. Operational Validation – Scored on an ordinal scale of 0-2 and then multiplied by 2. The degree of realism and adherence to military doctrine that the behavior's operational outputs yield. This measures how well the outputs represent the real life system or behaviors (Sargent, 2011).

3. Documented – Scored on an ordinal scale of 0-3. The volume and robustness of documentation that catalogs the behavior's conceptual model design, computer model verification, operational validity and data validity (Sargent, 2011). There are three distinct categories of documentation that are described in the Implementation and Technical Considerations portion of the meta-model that are in current use by TRAC WSMR.

4. SME – Scored on an ordinal scale of 0-2. The credibility of the subject matter expert who helped design the behavior.

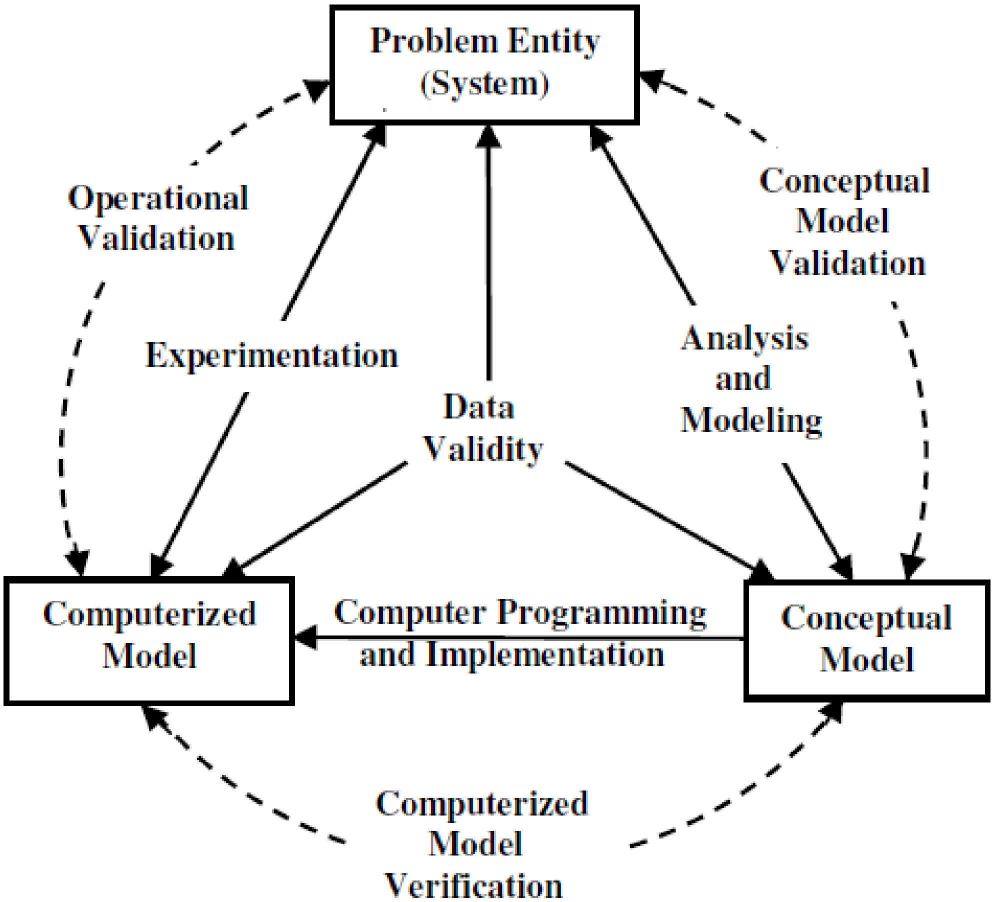


Figure 5. Simplified Version of the Modeling Process from (Sargent, 2011).

This scoring model has the capability of synthesizing the validation information associated with a model behavior and assigning it a score or degree of validation. The advantage of the final model is that components that speak to the model's realism (Conceptual Validation, Operational Validation, SME) account for 77% of the potential score, while the academic pedigree (Documented) accounts for only 23% of the potential score. We felt that this breakdown of the validation score put the emphasis where it should be and would better speak to the validity of behaviors. The final validation score tiers (Red, Amber, Green) also served to prioritize which behaviors needed the most refinement and which ones were the most sufficient for reuse in future COMBATXXI scenario builds. As a proposed practice we believe that the

‘Red’ tier indicates behaviors that are severely underdeveloped and need further development before they should be used in scenario development. The ‘Amber’ tier are behaviors that are generally sufficient for use but still could use further development to improve confidence in the results of a simulation model. The ‘Green’ tier would be behaviors that are highly preferable for scenario development and do not require any additional development if resources are limited.

Range	Criteria	Description	Score Assignment
0 - 2 (x2)	Conceptual Validation	Conceptual model supported by doctrine and is realistic.	Subjective score: 0-Poor, 1-Good, 2-Excellent
0 - 2 (x2)	Operational Validation	Behavior outputs are supported by doctrine and realistic.	Subjective score: 0-Poor, 1-Good, 2-Excellent
0 - 2	SME	Behavior was conceptualized by a credible SME.	SME level of expertise. 0-Poor, 1-Good, 2-Excellent.
0 - 3	Documented	Documentation: Conceptual Model. Confirmation: Behavior works correctly. Verification: Realism of outputs.	Level of Documentation: 0-None, 1-1/3 Doc types, 2-2/3 Doc types, 3-3/3 Doc types.

Validation Scale		
0 – 5 Strongly recommend further development.	**6 – 10** Recommend further development.	**11 – 13** Does not require further development.

Figure 6. Final Validation Scoring Model.

3.2.4.3. Validation Scoring Process for Standard Combat Behavior

Our validation scoring process is designed to work by allowing the raters/validators to use the content in the meta-model to guide them to score each one of the validation components in the scoring model. First, a rater reviews the meta-model, paying special attention to the ‘Validation Considerations’. Within the validation considerations the meta-data states will state if the behavior is combat behavior, indicated by a warfighting function, or a human behavior, indicated by the human behavior components. These two portions of the meta-data provide the ‘lens’ that the rater will evaluate the behavior. Once, the ‘lens’ is determined then the rater scores the Conceptual Validation and Operational Validation portions of the meta-model using one of the traditional validation techniques that are also contained in the Validation Considerations portion of the meta-model.

In use case scenario 2, we evaluated a Unit behavior called ‘SquadMove’, which involved a squad’s movement in an urban environment, see 4.3. Use Case scenario 2. If there was no buildings near the squad it would move in a file along the route, but if there were buildings close to the squad it would divide the force in parallel, assign team leaders and begin a bounding overwatch maneuver from building to building on opposite sides of the route. This behavior was classified as a ‘Maneuver’ warfighting function, which provided the lens that the rater would then score it. Based off of maneuver doctrine, referenced from FM3-06 Urban Ops and FM3-21.8 Infantry Rifle PLT/SQD we applied the validation technique face validation to the Conceptual Validation criteria of the scoring model. Based off of my combat experience and the doctrinal references, I assigned a score of 1 (Good) to the conceptual model of the behavior because it followed doctrine and my military experience of how that unit should maneuver. It would have received a 2 (Excellent), but I noticed that the conceptual model did not address highly unbalanced and unsymmetrical urban build up which would typically force an infantry squad to maneuver through the danger area in an alternative bounding overwatch maneuver used when crossing a linear danger area due to the squad’s exposed flanks (FM 3-21.8, 2007, pp. 3-35). Next, I ran two test scenarios, one with no near buildings and another with near buildings. In both scenarios the squad successfully maneuvered as the behavior was designed. Based off of this I applied the ‘Traces’ validation technique to these outputs at determined that the Operational Validation criteria scored a 2 (Excellent) because in the test scenario and given the validation technique applied the squad performed in accordance with doctrine and was reasonably realistic. I applied this score with some misgivings because the test scenarios did not address my before mentioned unbalanced urban terrain concern, but given the information I had I still applied the score of 2. Next, I addressed the SME rating in the scoring model. The documentation and meta-data stated that this behavior was created with the input of USMC and NPS MOVES Faculty SMEs. This was vague, but it did indicate that there were combat officers involved in its creation and since I was rating this behavior and it satisfied my personal doctrinal and combat experience I determined to apply a score of 1 (Good) to the credibility of the SME. If I had more specifics on the actual knowledge or position of the involved SME I may have assigned a score of 2. Last, I looked at the available documentation for this behavior. I had access to only one document series of documents all on implementation of the behavior and its conceptual model. This fit the "Documentation" category under the 'Documented' criteria. Since there was no Confirmation or Verification documents, I applied a 1 (1/3 Document Types) to the 'Documented' criteria. The final score for this behavior is shown in Table 1.

Conceptual Validation (x2)	2 (1 x 2 = 2)	Good: Realistic Soldier decision making process tied to doctrine. Does not address unbalanced urban build up.
Operational Validation (x2)	4 (2 x 2 = 4)	Excellent: Both squads performed as doctrinally expected.
SME Rating	1	Good: NPS Faculty w/ USMC advisors: Exact SME not listed but the collaboration and doctrinal references are acceptable.
Documented	1	Documentation that addresses the HTN conceptual model is present. Confirmation and Verification documents not found.
Validation Score	8	Recommend further development.

Table 1. Notional 'SquadMove" Validation Score.

3.2.4.4. Validation Scoring Process for Human Behavior

In some cases the M&S community may desire to extend behavior validation and documentation to account for the complexities of the human condition. The warfighting functions apply to how soldiers and units should behave in a combat scenario, but not all model elements are soldiers (e.g., civilians on the battlefield) and even soldiers are subject to human considerations. In order to capture the complexity of behaviors that are distinctly human it is important to consider behavior from a psychology perspective as well. Without this perspective many of the complexities and limitations of human behavior can be missed. In short a psychology perspective of human behavior will help define how a human will actually behave while attempting to adhere to the warfighting functions. Table 2 is a proposed human behavior model, consisting of six human components used to model agents in a simulated environment (Silverman, Bharathy, O'Brien, & Cornwell, 2006). By expanding our perspective of agent behavior we are able to analyze behavior at a human level and extend validation to the complexities of human psychology and non-combatant human agents.

Affect	Emotion or attitude generated from stimuli.
Perception	Organization, identification, and interpretation of sensory information in order to represent and understand the environment.
Cognition	Mental process related to problem solving and decision making.
Biology	Physical and mental capabilities and limitations (e.g., effects of stress and fatigue on Soldier, weight and speed limitations).
Social	Effects of social norms and pressures on human behavior.
Memory	Process of storing and recalling information.

Table 2. Human Behavior Components

In practice, scoring the validity of human behaviors was more complicated than scoring combat behaviors. We found that while some behaviors could be purely human, in COMBATXXI most were a combination of human and combat. Due to the limited quantity of unclassified behaviors we primarily identified applicable human behaviors that were also combat behaviors. If the desire is to take into account the human dynamic when validating a combat behavior we found that the validator took on the added onus of looking at the behavior through the lens of its applicable human components as well as doctrine. The result was that a combat behavior that would otherwise be a high scoring behavior because it performed as it doctrinally should could potentially score lower because human limitations could possibly create less satisfactory results. Although we did not find a clear example of this in COMBATXXI, you can imagine that human limitations could induce a fog-of-war effect on units and thus make it perform in a manner that was not consistent with how they should perform according to doctrine. An additional complexity that human behaviors raised was that unlike combat behaviors that generally could be attributed to just one or two warfighting functions, human behaviors tend to have multiple applicable human components. This increases the necessity for an analyst rating that behavior to evaluate all the applicable human components and adjudicate the best overall score in consideration to the multiple components.

Although the purpose and nature of physics-based and human behavior models are different, the validation processes are much the same (Goerger, McGinnis, & Darken, A Validation Methodology for Human Behavior Representation Models, 2005). In order to score a human behavior, we approached the task in much of the same manner that we would approach a pure combat behavior. The main difference was that we would take into account the six human components that are defined in the Validation Considerations portion of the meta-model. These components would act as the lens that we would measure the behavior validity. The only criteria of our scoring model significantly affected by validating a behavior by its human components are the conceptual and operational validation scores. The other two criteria, SME and Documented are unchanged.

A notional example could be applied to the 'SquadMove' example discussed earlier. In addition to the maneuver warfighting function we could say that the human components of 'perception' and 'cognition' are applicable to this behavior. Perception is applied by the ability of the squad leader to identify correctly the nearness of buildings to his route using his eyesight. Cognition is the squad leader's ability to process his tactical situation and make the decision to conduct bounding overwatch and determine the length of each bound. With these human components identified the validator would then apply traditional validation techniques to each one of these components. For the conceptual validation score I would likely apply face validation to state that the behavior does not address the ability of the squad leader to see his surroundings because the behavior automatically calculates the exact distance to each building, with no error, giving the squad leader perfect perception. Although humans rarely have perfect perception, this analyst believes the task of judging whether or not buildings are relatively close to you is a simple task, so the perfect perception in this case is not a huge deterrent.

Similarly, I could also apply face validation to the team leader's cognition and assess that it is reasonable that the squad leader could understand his tactical decision and based off his training make the correct doctrinal decision to implement the correct form of maneuver along his route. If the analyst has a stricter demand to more rigorously address this human component then he could observe that the 'SquadMove' behavior does not address different levels of training and combat experience, which could make different squad leaders make different decisions. Furthermore, the behavior still does not address uneven urban build up which the squad leader's training should have allowed him to make alternate maneuver decisions. Based off of this quick analysis I could give a conceptual validation score of 1 (Good) because the conceptual model looks satisfactory from a human component perspective, but lacks some of the complexity that would possibly produce unexpected results typical in daily human actions.

Operational validation can be approached the same way as we approached conceptual validation. The outputs of the behavior can be viewed from the perspective of how the human components should have behaved. For instance, an operational validation score of 1 (Good) could be applied to this behavior because using the 'traces' validation technique shows that while the squad maneuvered in the correct doctrinal manner in both test cases (buildings near and far), there may not have been any degree of variance in how they maneuvered. If this behavior was more robust and incorporated more complex human perception and cognition algorithms then we could expect to see a greater degree of variance in how the unit maneuvered.

After we weight the above ratings for the conceptual and operational validation scores and combine them with the previous SME and Documented scores, the final validation score is 6, see appendix b. Use Case Scenario 2. A score of 6 barely falls in our middle 'Amber' tier and suggests that it requires more development, although it is still useful for scenario development.

3.3. VALIDATING MODEL BEHAVIORS SUMMARY AND CONCLUSION

In this chapter we discussed the framework of our behavior validation process. This was approached from a knowledge management perspective designed to provide clarity and transparency to the diverse and interconnected nature of COMBATXXI behaviors, while delivering a validation score that would reflect the realism and integrity of those behaviors. This was done with a robust meta-model that contained behavior description data, validation considerations data, implementation and technical considerations, and a scoring model that provided a final validation score to model behaviors. It is important to note that our behavior validation process does not outline how to use the myriad of validation techniques used to score the conceptual and operational validation portions of the scoring model because the multitude of ways to implement these approaches to the huge quantity of diverse behaviors is too vast. However, we believe that this framework is a significant and crucial step in providing an overall validation process that can be applied to all behaviors.

SECTION 4. ANALYZING BEHAVIOR VALIDATION PROCESS

4.1. INTRODUCTION

Our behavior validation process consists is designed to add clarity and transparency to COMBATXXI behaviors and provide a final validation score to those behaviors. In order to validate our process, we investigated three use case scenarios to verify that our meta-model and scoring model could be used with various COMBATXXI behaviors. The three use case scenarios were unclassified COMBATXXI scenarios that had multiple behaviors associated with them. They were selected because they had a sufficient degree of documentation that allowed us to trace the role of each behavior in the scenario. From those use case scenarios we chose a sampling of the behaviors to enter into our meta-model as a proof of concept.

4.2. USE CASE SCENARIO 1 - UAS

The first use case scenario came from the Deployable Force Protection study by TRAC MTRY. This study analyzed the possible advantages of using a Boomerang threat detection capability to identify enemy indirect fire and automatically launch a UAS to search for the enemy indirect fire location and initiate accurate fires against it. As a base, the scenario used the 'Infantry at Risk – Vignette 3, Combat Outpost (COP) Attack' that simulates an indirect fire attack against an American combat outpost with organic counter fire capabilities and UAS assets. In the scenario a U.S. COP is attacked by a threat mortar system. The Boomerang locates the source and automatically launches a UAS. The UAS travels to threat quadrant, detects the threat and U.S. counter fire is initiated, see Figure 7.

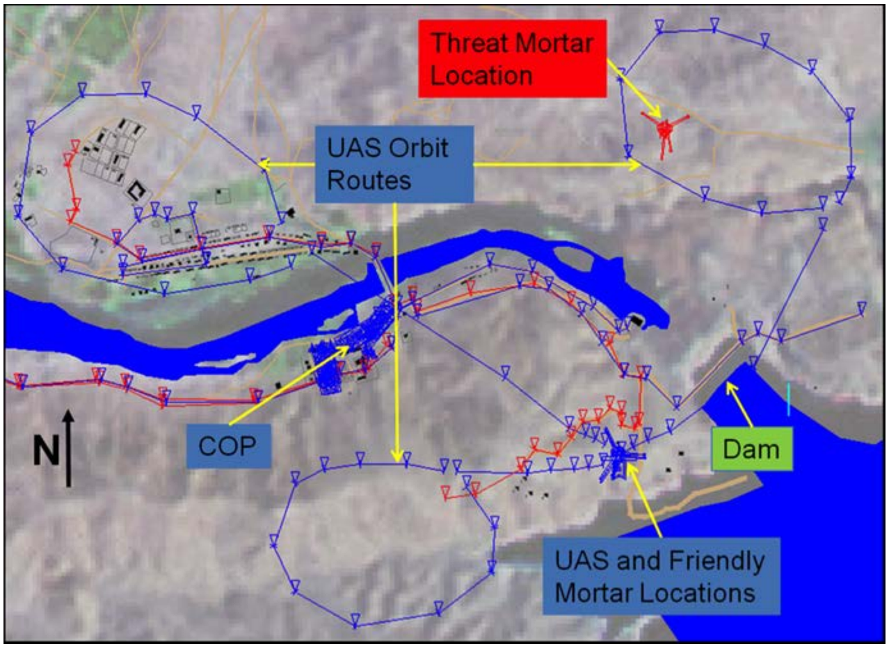


Figure 7. Use Case Scenario 1 Graphics.

The scenario is composed of two behavior packages:

1. Initialization – Necessary behaviors that initialize the starting conditions of the entities (e.g., permits units to fire if ordered to, establishes maneuver routes).

2. Default – A collection of 13 active behaviors that are interconnected and make red and blue forces initiate and react to the combat outpost attack. For the purpose of this scenario, it is beneficial to distinguish between red and blue forces. I have defined a sub-package behavior as 'UASAutoLaunch', which includes eight relevant and connected blue force behaviors that facilitate the launch of the UAS and the initiation of counterbattery fire. The remaining behaviors are red behaviors that allow them to initiate fires against the COP.

The study team focused on the UASAutoLaunch behavior because it was a good example of a complex combat behavior with multiple child behaviors. We found that this behavior fit neatly within our meta-model. We were able to define and categorize this behavior in our validation framework and assign a validation score based off of doctrinal reference. We were further able to define and categorize the child behaviors using our meta-model, but many of the child behaviors were very simple procedural behavior types, with one or two triggers that resulted in a single action. These were more challenging to assign a validation score because they were very simple and did not seem realistic without the context of the greater parent behavior. The way we approached these scores was to merely assess its functionality within the greater UASAutoLaunch behavior and assign it a high conceptual and operational validation score if it was a reasonable behavior that allowed the parent behavior to work properly. Another possibility that we considered was not to score it at all, and only score the complex parent behaviors because they represented the full combat behavior. We decided against this approach because for our immediate purposes it was beneficial to evaluate each behavior in terms of validity despite the simplicity of procedural and primitive behaviors to ensure that we scrutinized all the components that built UASAutoLaunch, see appendix a. use case scenario 1.

4.3. USE CASE SCENARIO 2 – URBAN SQUAD MANEUVER

The second use case scenario that the study team analyzed was the Naval Postgraduate School (NPS) Bounding Overwatch HTN Tutorial. The scenario was developed by the NPS MOVES department with support from USMC. In this scenario two squads are conducting a patrol in an urban environment. One is on a street without buildings the other is on a highly built up street. Commanders are required to choose and execute the correct maneuver technique based on the terrain. The behavior's conceptual model is designed to maneuver the squad in one of two ways depending on how close buildings are along the squad's route. If buildings are within 10 meters of the squad's route the squad leader will split his forces into two bounding elements and begin bounding overwatch along the route until it reaches its destination or the urban terrain becomes less dense. If the buildings are further than 10 meters from the route, then the squad will travel along the route in a wedge formation, see Figure 8.

The dominant behavior is 'SquadMove', which is a complex behavior, developed in a newly developed hierarchal task network GUI that allows a squad to make a tactical decision on how to maneuver in urban terrain. This behavior has two child behaviors called InitGoals and JumpStart that are essentially primitive behaviors that allow the user to designate the start and end locations and then call on the SquadMove behavior to execute the maneuver sequence.

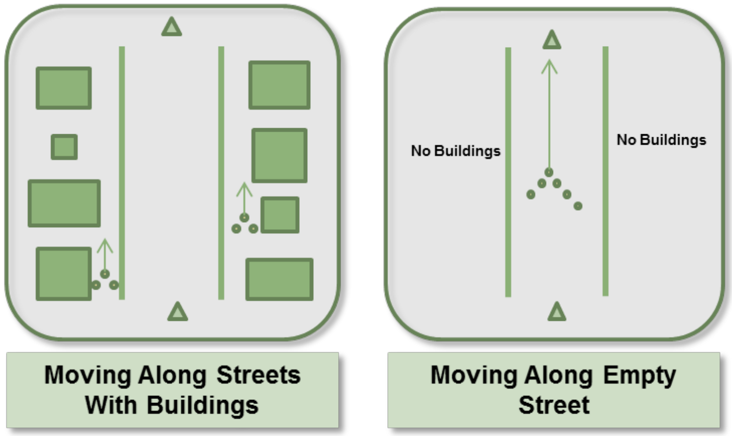


Figure 8. Illustration of Urban Maneuver Behavior.

We chose this scenario for two reasons. First, it was a scenario that contained a human behavior in addition to the obvious combat component. Second, it was created using non-traditional means using the hierarchal task network GUI. The results from applying our validation process to this scenario showed that our meta-model was sufficient for capturing human behaviors and allowed a validation score that addressed both the human and combat realism of the behavior. It also reconfirmed the difficulty of assigning validation scores to simple child behaviors because they have little realistic meaning on their own without direct connection to their parent behavior. Despite this difficulty, it was still possible to assign a score and the validation process did contribute to the transparency of the behaviors, reference appendix b. Use Case Scenario 2.

4.4. USE CASE SCENARIO 3 - BOCME

Use Case Scenario 3 is a tutorial called Basic Observer, Communication, Move, and Engage (BOCME) scenario. It is not designed to be a realistic scenario, populated with the most sophisticated behaviors, but it is a good example of how a series of human behaviors can make soldier entities perform a simple scenario. The BOCME scenario involves three entities (Red Soldier, Blue Soldier 1, Blue Soldier 2). The scenario involves eight primary behaviors that are tactical, procedural and primitive. There are also approximately 12 primitive component behaviors, referred to as actions in SITS, that allow the procedural and tactical behaviors to work. This scenario is loosely based on a point ambush where the Red Soldier begins movement on a route. Along that route is Blue Soldier 1 who is observing a section of the route from a hide position. Blue Soldier 2 is hidden from the route waiting on an observation report from Blue Soldier 1. The scenario begins with the Red Soldier executing movement along the prescribed route. When Blue Soldier 1 observes the Red Soldier on the route, he sends a radio message to Blue Soldier 2 who then moves to an ambush site at another location on the route. Once Blue Soldier 2 sees the Red Soldier, he engages with whatever weapon he has been assigned, see Figure 9 (BOCME Scenario Builders Guide, 2011). This is not a doctrinally realistic scenario but allows scrutiny of simple soldier behaviors that can govern a scenario.

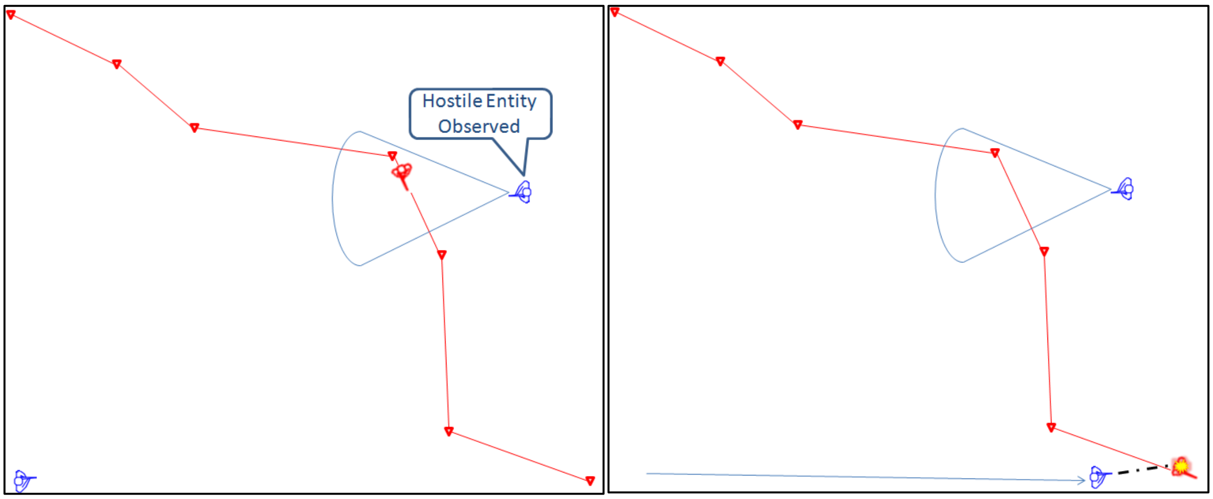


Figure 9. BOCME Scenario from (BOCME Scenario Builders Guide, 2011).

4.5. USE CASE SCENARIO SUMMARY AND CONCLUSION

In this chapter we discussed two use case scenarios that we used to test our validation process and meta-model framework. Scenario one provided an example of a complex combat behavior and its child behaviors in a COP attack and the corresponding unmanned aerial vehicle launch and counter fire response. Scenario two identified a complex human behavior and two primitive behaviors that controlled how a squad maneuvered tactically in an urban environment. In both use case scenarios we found that our meta-model was sufficient to define, describe, and classify all the behaviors and add transparency to the relationship of the behaviors. We also saw that the scoring model worked on the behaviors, but was more useful in scoring the validity of the complex behaviors and less useful when trying to score the simpler child behaviors. Despite this inherent difficulty, we made the decision to score the child behaviors in order to ensure that we analyzed each behavior for logical implementation and realism when applicable.

SECTION 5. SUMMARY

5.1. SUMMARY

The purpose this project was to develop a methodology to validate and document simulation model behaviors in COMBATXXI in order to increase transparency and improve the ability of users to identify and reuse valid behaviors in the creation of future COMBATXXI scenarios. This study team approached this from a knowledge management perspective by creating a validation process that described and defined behaviors and their relationships to other behaviors and scenarios and provided a scoring model that would rate behaviors on their degree of validity. This report outlined the meta-model framework developed to describe, define, categorize and finally score model behaviors. This meta-model had four broad categories: 1) Behavior Description Data – Generalized data that classifies and describes the behavior. 2) Validation Considerations – Information that was considered in the validation of each behavior. 3) Implementation and Technical Considerations – Information and documentation critical to understanding how the behavior works, how it has been used and how it is implemented. 4) Validation Score – A validation score computed by a scoring model developed by this study team to rate the overall level of validity that a behavior demonstrated. This scoring model was further demonstrated using an urban maneuver behavior that was a component in a use case scenario. Finally, this report described two use case scenarios that this study team used to validate our meta-model and scoring model. This showed that our methodology was sufficient, but there were inherent difficulties in scoring the realism/validity of very simple child behaviors, because they had little meaning independent from their parent behavior.

5.2. CONCLUSIONS AND RECOMMENDATIONS

Validating COMBATXXI behaviors is a complex task due to the variety and volume of model behaviors. However, it is critical that a behavior validation process be established to introduce transparency to behavior quality and interconnectedness so that behaviors can be quickly reused in future scenarios and so that the integrity of study results can traced to the validity of their corresponding behaviors that led to those results. During the course of this project some key conclusions became apparent on the topic of our proposed validation methodology.

1. The meta-model must be robust enough to define, describe and categorize the enormous variety of COMBATXXI behaviors. We believe that our meta-model is sufficiently robust to capture the different parent and child behaviors contained in COMBATXXI. After analyzing multiple behaviors in two use case scenarios we believe that our meta-model will correctly categorize all types of behaviors, including human behaviors, and add transparency to the transparency of COMBATXXI behaviors.
2. The scoring model validation approach is an effective way to validate behaviors. It provides a tool that can be applied by numerous individuals and places strong emphasis on the behavior's realism by scoring the conceptual and operational validation. This is further reinforced by the quality of the SME who adds their experience and knowledge to the development of that behavior.
3. Complex parent behaviors are more easily validated than their simpler child behaviors by our scoring model. This is because the parent behavior is generally a complete process and more relatable to human experience, while the child behaviors are often so simplistic that they have little meaning when viewed independent of their parent. We recommend evaluating each child behavior in context with its parent behavior and score it high if it facilitates the successful implementation of the parent behavior. This will ensure that each child behavior is addressed. In later refinement, an organization may decide not to apply the scoring model to simplistic child behaviors if it does not seem to apply (arguably with primitive and procedural behaviors), but this refinement should not be applied until a trial phase has already been implemented.
4. Due to the large volume of behaviors it is unreasonable to ask a select group of evaluators to validate each behavior. We recommend that the behavior developer with the support of the SME assign the validation score. This score then would be accepted or denied by an administrative authority within the COMBATXXI community. This allows for a divided effort to validate behaviors by individuals with intrinsic understanding of how the behavior was created and how it performs.

5. Validators need leeway to choose which validation technique to employ when validating a behavior. The large variety of behaviors requires complete freedom to select the most applicable validation technique. However, traditional validation techniques must be employed to make the validation score academically relevant. Arbitrary validation scores would undermine the entire process. This is why documentation must support each behavior and why it is a part of the scoring model.

6. Validating human behaviors is more complicated than validating combat behaviors because human behaviors have multiple components at play. Additionally, human components often have to be applied in addition to the doctrinal warfighting functions critical to evaluating a behavior that is both combat and human. The scoring model allows both types of behaviors to be scored; it only requires that the validator views the conceptual and operational validation through the lens of the human components and if necessary weigh it next to the behavior's adherence to doctrine. However, the application of a traditional validation technique remains a requirement when applying a validation score.

7. We recommend that our validation process should be initially reviewed and refined by TRAC WSMR. The final product then should be implemented in a trial run as the COMBATXXI library. During that trial period refinement to the meta-model can be done as irrelevant and relevant meta-data is identified by users. Additionally, further refinement to the scoring model can be conducted as more users and administrators review validation scores. It is foreseeable that the scale and weights of each criterion could be improved upon in order to make scoring easier for the validators and more operationally relevant to the COMBATXXI community.

8. We recommend that military modeling and simulation communities strive to make behavior libraries network accessible. This would allow multiple organizations to collaborate and reuse well developed behaviors. It would also provide greater transparency across the community. In Appendix E, we describe a proof-of-concept implementation of a network accessible library using available and secure network technologies.

APPENDIX A. USE CASE SCENARIO 1 - UAS

Use case scenario 1 captures a complex combat behavior that initiates when a COP is attacked by enemy mortar fire and ends once a UAS is launched and the enemy mortar system is detected and destroyed.   This behavior consists of nine component behaviors that work together to launch the UAS, search and find the enemy mortar system, and initiate counter battery fire. The below tables show how we applied the meta-model to a sampling of the behaviors.

Behavior Name	UASAutoLaunch	mnvr_pass_UAS_order	Permit_Engagement	Blue_Engage
Date	15-Sep-12	15-Sep-12	15-Sep-12	15-Sep-12
Behavior Description (to include trigger and action)	Deployable Force protection capability to auto launch a UAS for targeting of enemy mortar. The acoustic direction finding was 0 / 1 capability based on whether or not the technology was present.	Gives order to launch UAS after detection of enemy mortar fire.	Necessary condition that allows a unit to fire rounds at a target. Generally addressed in initialization behaviors.	When UAS detects enemy mortar it creates a potential target list and gives command to Blue_Mortar to engage target.
Parent Behavior	N/A	UASAutoLaunch	UASAutoLaunch BlueShootArtillery_49	UASAutoLaunch
Component Behaviors	Blue_Engage, BlueBeginAerialSurveil BlueShootArtillery_49 mnvr_exec_order, mnvr_pass_RED49_loc mnvr_pass_UAS_order Permit_Engagement mnvr_startdelay mnvr_trigger_mgr			BlueShootArtillary_49, Permit_Engagement
Associated Vignettes	Infantry at Risk – Vignette 3 Combat Outpost (COP) Attack	Infantry at Risk – Vignette 3 Combat Outpost (COP) Attack	Infantry at Risk – Vignette 3 Combat Outpost (COP) Attack	Infantry at Risk – Vignette 3 Combat Outpost (COP) Attack
Key Entities	Blue_Mortar, Red_Mortar, Blue_UAS, Blue_COP	Blue_Mortar, Red_Mortar, Blue_UAS, Blue_COP	Blue_Mortar,	Blue_Mortar, Blue_UAS,
Resolution	( ) >Subsystem	( ) >Subsystem	( ) >Subsystem	( ) >Subsystem
	( ) >Entity	(X) >Entity	(X) >Entity	(X) >Entity
	( ) >Unit	( ) >Unit	( ) >Unit	( ) >Unit
	(X) >Operational	( ) >Operational	( ) >Operational	( ) >Operational
Behavior Type	(X) >Cognitive	( ) >Cognitive	( ) >Cognitive	( ) >Cognitive
	( ) >Tactical	( ) >Tactical	( ) >Tactical	(X) >Tactical
	( ) >Procedural	(X) >Procedural	( ) >Procedural	( ) >Procedural
	( ) >Primitive	( ) >Primitive	(X) >Primitive	( ) >Primitive

Table A.1. UASAutoLaunch Behavior Description Data.

Behavior Name	UASAutoLaunch	mnvr_pass_UAS_order	Permit_Engagement	Blue_Engage
Validation Technique	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition (X) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition (X) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition (X) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition (X) >Traces
Warfighter Function	( ) >Maneuver ( ) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment (X) >Fires (X) >Protection	( ) >Maneuver ( ) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires (X) >Protection	( ) >Maneuver ( ) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment (X) >Fires ( ) >Protection	( ) >Maneuver ( ) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment (X) >Fires ( ) >Protection
Human Behavior (Y/N)	N	N	N	N
Human Component (select and specify degree accounted for low/med/high)	( ) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)
Reference Doctrine	FM 3-04.155 UAV Operations	FM 3-04.155 UAV Operations	FM 3-04.155 UAV Operations	FM 3-04.155 UAV Operations
SME	TRAC Officer: MAJ Ed Masotti and MAJ Pete Nesbitt	TRAC Officer: MAJ Ed Masotti and MAJ Pete Nesbitt	TRAC Officer: MAJ Ed Masotti and MAJ Pete Nesbitt	TRAC Officer: MAJ Ed Masotti and MAJ Pete Nesbitt

Table A.2. UASAutoLaunch Validation Considerations.

Behavior Name	UASAutoLaunch	mnvr_pass_UAS_order	Permit_Engagement	Blue_Engage
Implementation Method	BSL, Python	BSL	BSL	BSL
Project Study	TRAC MTRY ASAALT DFP	TRAC MTRY ASAALT DFP	TRAC MTRY ASAALT DFP	TRAC MTRY ASAALT DFP
Reference Model Implementation	COMBATXXI	COMBATXXI	COMBATXXI	COMBATXXI
Reference Model Version	Stablebuild_20130912	Stablebuild_20130912	Stablebuild_20130912	Stablebuild_20130912
Reference Documentation	TRAC-M-TR-13-019 January 2013, COMBATXXI Modeling and Simulation Methodology for New Deployable Force Protection Technology	TRAC-M-TR-13-019 January 2013, COMBATXXI Modeling and Simulation Methodology for New Deployable Force Protection Technology	TRAC-M-TR-13-019 January 2013, COMBATXXI Modeling and Simulation Methodology for New Deployable Force Protection Technology	TRAC-M-TR-13-019 January 2013, COMBATXXI Modeling and Simulation Methodology for New Deployable Force Protection Technology
POC	Dr. Thomas Anderson (ERDC CRREL)	Dr. Thomas Anderson (ERDC CRREL)	Dr. Thomas Anderson (ERDC CRREL)	Dr. Thomas Anderson (ERDC CRREL)
Security Distribution	Unclassified	Unclassified	Unclassified	Unclassified
Distribution	Unlimited	Unlimited	Unlimited	Unlimited

Table A.3. UASAutoLaunch Implementation and Technical Considerations.

Behavior Name	UASAutoLaunch	mnvr_pass_UAS_order	Permit_Engagement	Blue_Engage
Conceptual Validation (0,2,4)	2-Behavior framework was reasonable and supported by doctrine, but the detection capabilities were too accurate to be 100% realistic.	4-Very reasonable conceptual model for soldiers to launch UAS after mortar attack.	4-Necessary primitive behavior that equates to a command giving its units the authorization to engage hostile forces. Implementation is very accurate.	2-Tactical behavior was reasonable and supported by doctrine. It was unrealistic that the UAS passed a 100% accurate grid locations and would accurately identify the target 100% of the time.
Operational Validation (0,2,4)	2-Results from engagement were reasonable, but detection capability and ability to pass accurate enemy location were unrealistic.	4-Very realistic launch results after attack initiated on COP.	4-Behavior performed function in scenario as designed and was consistent with units that are authorized to engage hostile targets.	2-Reasonable results but questionable because of the assumptions in the conceptual model.
SME Developed (0-2)	2-SMEs were very knowledgeable with company level combat experience with similar operations.	2-SMEs were very knowledgeable with company level combat experience with similar operations.	2-SMEs were very knowledgeable with company level combat experience with similar operations.	2-SMEs were very knowledgeable with company level combat experience with similar operations.
Documented (0-3)	3-Documentation, Verification and Confirmation supporting documents were included in the TRAC Report	1-Low complexity behaviors were not fully supported in TRAC report. Documentation was located in the COMBATXXI code.	1-Low complexity behaviors were not fully supported in TRAC report. Documentation was located in the COMBATXXI code.	1-Low complexity behaviors were not fully supported in TRAC report. Documentation was located in the COMBATXXI code.
Total Score (0-13)	9-Recommend further development.	11-Does not require further development.	11-Does not require further development.	7-Recommend further development.

Table A.4. UASAutoLaunch Validation Scores.

APPENDIX B. USE CASE SCENARIO 2 – SQUAD MANEUVER

Use case scenario 2 captures a complex behavior called SquadMove designed to control the maneuver of an infantry squad in an urban environment. This behavior consists of one parent behavior and two component behaviors that work together to maneuver an infantry squad along a route in an urban environment. This behavior is unique because it was created using a HTN GUI that allows that squad to make decisions on their environment and make tactical decisions based off of their situational awareness. The below tables show how we applied the meta-model to a sampling of the behaviors.

Behavior Name	SquadMove	InitGoals	JumpStart
Date	15-Nov-12	15-Nov-12	15-Nov-12
Behavior Description (to include trigger and action)	Platoon maneuver in an urban setting. Unit implements correct maneuver technique based on urban terrain. Bounding Overwatch vs. Wedge Formation.	Behavior allows the user to designate the start and end points of the route.	Behavior calls on the SquadMove behavior to initiate the maneuver.
Parent Behavior	N/A	SquadMove	SquadMove
Component Behaviors	InitGoals JumpStart	N/A	N/A
Associated Vignettes	Urban Raid Scenario, NPS Bounding Overwatch HTN Tutorial	Urban Raid Scenario, NPS Bounding Overwatch HTN Tutorial	Urban Raid Scenario, NPS Bounding Overwatch HTN Tutorial
Key Entities	Squad Soldier	Squad Soldier	Squad Soldier
Resolution	( ) >Subsystem	( ) >Subsystem	( ) >Subsystem
	( ) >Entity	( ) >Entity	( ) >Entity
	(X) >Unit	(X) >Unit	(X) >Unit
	( ) >Operational	( ) >Operational	( ) >Operational
Behavior Type	(X) >Cognitive	( ) >Cognitive	( ) >Cognitive
	( ) >Tactical	( ) >Tactical	( ) >Tactical
	( ) >Procedural	( ) >Procedural	( ) >Procedural
	( ) >Primitive	(X) >Primitive	(X) >Primitive

Table B.1. SqaudMove Behavior Description Data.

Behavior Name	SquadMove	InitGoals	JumpStart
Validation Technique	(X) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces
Warfighter Function	(X) >Maneuver (X) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection	(X) >Maneuver (X) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection	(X) >Maneuver (X) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection
Human Behavior (Y/N)	Y	N	N
Human Component (select and specify degree accounted for low/med/high)	(L) >Perception (H/M/L) (M) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)
Reference Doctrine	FM3-06 Urban Ops, FM3-21.8 Inf Rifle PLT / SQD, FM3-24 COIN	FM3-06 Urban Ops, FM3-21.8 Inf Rifle PLT / SQD, FM3-24 COIN	FM3-06 Urban Ops, FM3-21.8 Inf Rifle PLT / SQD, FM3-24 COIN
SME	USMC / NPS MOVES Faculty	USMC / NPS MOVES Faculty	USMC / NPS MOVES Faculty

Table B.2. SquadMove Validation Considerations.

Behavior Name	SquadMove	InitGoals	JumpStart
Implementation Method	HTN, Python	HTN, Python	HTN, Python
Project Study	Marine Personnel Carrier (MPC) variants in an urban environment.	Marine Personnel Carrier (MPC) variants in an urban environment.	Marine Personnel Carrier (MPC) variants in an urban environment.
Reference Model Implementation	COMBATXXI	COMBATXXI	COMBATXXI
Reference Model Version	V2.3 (2013-2014)	V2.3 (2013-2014)	V2.3 (2013-2014)
Reference Documentation	Using Hierarchical Task Networks in COMBATXXI: Bounding Overwatch HTN Tutorial	Using Hierarchical Task Networks in COMBATXXI: Bounding Overwatch HTN Tutorial	Using Hierarchical Task Networks in COMBATXXI: Bounding Overwatch HTN Tutorial
POC	Dr. Imre Balough	Dr. Imre Balough	Dr. Imre Balough
Security Distribution	Unclassified	Unclassified	Unclassified
Distribution	Unlimited	Unlimited	Unlimited

Table B.3. SquadMove Implementation and Technical Considerations.

Behavior Name	SquadMove	InitGoals	JumpStart
Conceptual Validation (0,2,4)	2-Doctrinally realistic, but does not address asymmetric urban build up and does not introduce complex cognition and perception components that would lead to greater variation in maneuver decisions.	4-Very reasonable conceptual model for soldiers to launch UAS after mortar attack.	4-Necessary primitive behavior that equates to a command giving its units the authorization to engage hostile forces. Implementation is very accurate.
Operational Validation (0,2,4)	2-Squads performed as doctrinally expected in limited use case scenario but did not introduce variation to decision making process or test their ability to maneuver in asymmetric urban environment.	4-Very realistic launch results after attack initiated on COP.	4-Behavior performed function in scenario as designed and was consistent with units that are authorized to engage hostile targets.
SME Developed (0-2)	1-Unspecified USMC contributors suggest a credible SME.	2-MOVES faculty is expert SME for the creation of this primitive behavior.	2- MOVES faculty is expert SME for the creation of this primitive behavior.
Documented (0-3)	1- Documentation that addresses the HTN conceptual model is present. Confirmation and Verification documents not found.	1- Documentation that addresses the HTN conceptual model is present. Confirmation and Verification documents not found.	1- Documentation that addresses the HTN conceptual model is present. Confirmation and Verification documents not found.
Total Score (0-13)	6-Recommend further development.	11-Does not require further development.	11-Does not require further development.

Table B.4. SquadMove Validation Scores.

APPENDIX C. USE CASE SCENARIO 3 – BOCME

Use case scenario 3 is a tutorial called Basic Observer, Communication, Move, and Engage (BOCME) scenario. It is not designed to be the most realistic scenario, populated with the most sophisticated behaviors, but it is a good example of how a series of human behaviors can make soldier entities perform a simple scenario. The BOCME scenario involves three entities (Red Soldier, Blue Soldier 1, Blue Soldier 2). The scenario involves eight primary behaviors that are tactical, procedural and primitive. There are also approximately 12 primitive component behaviors (referred to as actions in SITS) that allow the procedural and tactical behaviors to work. This scenario is loosely based on a point ambush where the Red Soldier begins movement on a route. Along that route is Blue Soldier 1 who is observing a section of the route from a hide position. Blue Soldier 2 is hidden from the route waiting on an observation report from Blue Soldier 1. The scenario begins with the Red Soldier executing movement along the prescribed route. When Blue Soldier 1 observes the Red Soldier on the route, he sends a radio message to Blue Soldier 2 who then moves to an ambush site at another location on the route. Once Blue Soldier 2 sees the Red Soldier, he engages with whatever weapon he has been assigned (BOCME Scenario Builders Guide, 2011). This is not a doctrinally realistic scenario but allows scrutiny of simple soldier behaviors that can govern a scenario.

Behavior Name	Observe	Communicate	MoveOnRoute	Engage
Date	2011	2011	2011	2011
Behavior Description (to include trigger and action)	Behavior makes a Soldier observe an area with his assigned sensor (eyes, binos, etc.) in the direction he is oriented. Trigger: AllEntitiesHaveBeenInitiated.	Sends message that threat has been seen. Trigger: TargetSeen	Calls on and executes compound order called CO_ROUTE, which moves entity along route waypoints. Trigger: AllEntitiesHaveBeen Initiated.	Allows entity to engage a target. Trigger: PotentialTargetList Complete.
Parent Behavior	N/A	N/A	N/A	N/A
Component Behaviors	ObserveSectorWithSensor	SetMsgFunction, SendMessage, AddMsgLine	processorder "CO_ROUTE"	PassPotentialTargetsToEngage
Associated Vignettes	BOCME	BOCME	BOCME	BOCME
Key Entities	Blue Soldier 1	Blue Soldier 1	Red Soldier	Blue Soldier 2
Resolution	( ) >Subsystem	( ) >Subsystem	( ) >Subsystem	( ) >Subsystem
	(X) >Entity	(X) >Entity	(X) >Entity	(X) >Entity
	( ) >Unit	( ) >Unit	( ) >Unit	( ) >Unit
	( ) >Operational	( ) >Operational	( ) >Operational	( ) >Operational
Behavior Type	( ) >Cognitive	( ) >Cognitive	( ) >Cognitive	( ) >Cognitive
	( ) >Tactical	(X) >Tactical	( ) >Tactical	( ) >Tactical
	(X) >Procedural	( ) >Procedural	(X) >Procedural	(X) >Procedural
	( ) >Primitive	( ) >Primitive	( ) >Primitive	( ) >Primitive

Table C.1. BOCME Behavior Description Data.

Behavior Name	Observe	Communicate	MoveOnRoute	Engage
Validation Technique	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces	( ) >Animation: ( ) >Comparison ( ) >Event Validity (X) >Face Validity ( ) >Extreme Condition ( ) >Traces
Warfighter Function	( ) >Maneuver ( ) >Mission Cmd ( ) >Communication (X) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection	( ) >Maneuver ( ) >Mission Cmd (X) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection	(X) >Maneuver ( ) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection	(X) >Maneuver ( ) >Mission Cmd ( ) >Communication ( ) >Intelligence ( ) >Sustainment ( ) >Fires ( ) >Protection
Human Behavior (Y/N)	Y	Y	Y	Y
Human Component (select and specify degree accounted for low/med/high)	(L) >Perception (H/M/L) ( ) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) (L) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	( ) >Perception (H/M/L) (L) >Cognition (H/M/L) (L) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)	(L) >Perception (H/M/L) (L) >Cognition (H/M/L) ( ) >Biology (H/M/L) ( ) >Memory (H/M/L) ( ) >Affect (H/M/L) ( ) >Social (H/M/L)
Reference Doctrine	FM3-21.8 Inf Rifle PLT / SQD	FM3-21.8 Inf Rifle PLT / SQD	FM3-21.8 Inf Rifle PLT / SQD	FM3-21.8 Inf Rifle PLT / SQD
SME	TRAC Officer: MAJ Adam Haupt	TRAC Officer: MAJ Adam Haupt	TRAC Officer: MAJ Adam Haupt	TRAC Officer: MAJ Adam Haupt

Table C.2. UASAutoLaunch Validation Considerations.

Behavior Name	Observe	Communicate	MoveOnRoute	Engage
Implementation Method	BSL	BSL	BSL	BSL
Project Study	TRAC WSMR Tutorial	TRAC WSMR Tutorial	TRAC WSMR Tutorial	TRAC WSMR Tutorial
Reference Model Implementation	COMBATXXI	COMBATXXI	COMBATXXI	COMBATXXI
Reference Model Version	Unknown	Unknown	Unknown	Unknown
Reference Documentation	BOCME Scenario Builders Guide	BOCME Scenario Builders Guide	BOCME Scenario Builders Guide	BOCME Scenario Builders Guide
POC	TRAC WSMR	TRAC WSMR	TRAC WSMR	TRAC WSMR
Security Distribution	Unclassified	Unclassified	Unclassified	Unclassified
Distribution	Unlimited	Unlimited	Unlimited	Unlimited

Table C.3. BOCME Implementation and Technical Considerations.

Behavior Name	Observe	Communicate	MoveOnRoute	Engage
Conceptual Validation (0,2,4)	2-Behavior accounts for low level of human perception, but the underlying COMBATXXI LOS function ensures observer can only perceive things that are in his LOS. Once target enters LOS, automatic perception (100% accuracy).	4-Very reasonable conceptual model for soldier to send contact report when target is seen using available communications equipment. The Low level that cognition was modeled in this behavior does not deteriorate this simple process.	0-Calls on CO_ROUTE which moves a Soldier along a route consisting of seven waypoints. Low biology and cognition accountability because it does not account for soldier's ability to navigate, vary movement technique or account for Soldier speed variance based off of terrain, physical condition or load out.	2-Low cognition component accounted for. Engages as soon as target is seen. No other considerations accounted for. Still a reasonable algorithm for a trained Soldier with a clear mission and engagement criteria.
Operational Validation (0,2,4)	2-Spotted target in LOS. A more sophisticated human behavior would take into account camouflage, daylight and attention of observer (better SA).	4-Soldier delivered message which is consistent with historical experience of Soldiers observing an engagement area.	2-Soldier conducted route as commanded but did not show any variance that would be expected from a human.	2-Reasonable results because the COMBATXXI shooting algorithm is a stochastic process based off of Soldier hit probabilities that account for range and target type.
SME Developed (0-2)	1-SME was reasonably knowledgeable with human observation characteristics in a combat environment.	1-SME was reasonably knowledgeable with human cognitive process of sending a simple contact report over the radio.	1-SME was reasonably knowledgeable with human observation characteristics in a combat environment.	1-SME was reasonably knowledgeable with human observation characteristics in a combat environment.
Documented (0-3)	1-Implementation documentation in the BOCME guide explained conceptual model only. No confirmation or verification documentation.	1-Implementation documentation in the BOCME guide explained conceptual model only. No confirmation or verification documentation.	1-Implementation documentation in the BOCME guide explained conceptual model only. No confirmation or verification documentation.	1-Implementation documentation in the BOCME guide explained conceptual model only. No confirmation or verification documentation.
Total Score (0-13)	6-Recommend further development.	10- Recommend further development.	4-Strongly recommend further development.	6-Recommend further development.

Table C.4. BOCME Validation Scores.

APPENDIX D. VALIDATION TECHNIQUES

Validation techniques provide an objective and subjective approach to validation. Objective approaches, usually imply the use of statistical tests and procedures, while subjective approaches rely on graphical displays, intuition, opinions or subject matter expertise (Birta & Ozmizrak, 1996). Physics based models tend to lend themselves to objective approaches; whereas HBMs tend to require the application of subjective approaches because much of the human mind is completely or partially unobservable. Selection of an appropriate validation technique can be a considerable problem (Birta & Ozmizrak, 1996). Generally, the most common means of validating conceptual models is face validation using a SME or collection of SMEs. Operational validation can be supported by objective validation techniques because the model produces outputs that can be measured against the real live system or other validated behavior models. As mentioned earlier, human behavior models are more difficult because much of the inner workings of the brain are not observable. This means that cognitive models are also most commonly validated using face validation using SME elicitation (Goerger, McGinnis, & Darken, A Validation Methodology for Human Behavior Representation Models, 2005). No matter which validation technique is chosen, it is important that the validating agent carefully selects a technique or collection of techniques that are realistically supported by the nature of the behavior model. Below in is a comprehensive list of validation techniques that have been historically used and recommended by Dr. Robert G. Sargent, the former President of the INFORMS Simulation Society, see **Table D.**1.

Animation	Model's operational behavior is displayed graphically as the model moves through time, e.g., the movements of parts through a factory during a simulation run are shown graphically.
Comparison to Other Models	Various results (outputs) of the simulation model being validated are compared to results of other (valid) models. For example, simple cases of a simulation model are compared to known results of analytic models.
Degenerate Tests	The degeneracy of the model's behavior is tested by appropriate selection of values of the input and internal parameters, e.g. does the average number in the queue of a single server increase over time when the arrival rate is larger than the service rate?
Event Validity	The "events" of occurrences of the simulation model are compared to those of the real system to determine if they are similar. For example, compare the number of fires in a fire department simulation to the actual number of fires.
Extreme Condition Tests	The model structure and outputs should be plausible for any extreme and unlikely combination of levels of factors in the system. For example, if in-process inventories are zero, production output should usually be zero.
Face Validity	Individuals knowledgeable about the system are asked whether the model and/or its behavior are reasonable. For example, is the logic in the conceptual model correct and are the model's input-output relationships reasonable?
Historical Data Validation	If historical data exist, part of the data is used to build the model and the remaining data are used to determine (test) whether the model behaves as the system does.
Historical Methods	The three historical methods of validation are *rationalism*, *empiricism*, and *positive economics*. Rationalism requires that assumptions underlying a model be clearly stated and readily accepted. Logic deductions are used from these assumptions to develop the valid model. Empiricism requires every assumption and outcome to be empirically validated. Positive economics requires only that the model's outcome(s) be correct and is not concerned with a model's assumptions or structure.
Internal Validity	Several replications of a stochastic model are made to determine the amount of variability in the model. A large amount may cause the model's results to be questionable.
Multistage Validation	Combining the three historical methods of rationalism, empiricism, and positive economics into a multistage process of validation. This validation method consists of (1) developing the model's assumptions on theory, observations, and general knowledge, (2) validating the model's assumptions by empirically testing them, and (3) comparing the input-output relationships of the model to the real system.
Operational Graphics	Values of various performance measures, e.g., the percentage of servers busy, are shown graphically as the model runs through time; i.e., the behaviors of performance indicators are visually displayed to ensure they behave correctly.
Parameter Sensitivity	Consists of changing the values of the input and internal parameters of a model to determine the effect upon the model's behavior or output. The same relationships should occur in the model as in the real system.
Predictive Validation	The model is used to predict the system's behavior, and then comparisons are made between the system's behavior and the model's forecast to determine if they are the same. The system data may come from an operational system or a set of experiments.
Traces	The behaviors of different types of specific entities in the model are traced through the model to determine if the model's logic is correct.
Turing Tests	Individuals who are knowledgeable about the operations of the system being modeled are asked if they can discriminate between system and model outputs.

Table D.1. Validation Techniques after (Sargent, 2011)

APPENDIX E. BEHAVIOR LIBRARY IMPLEMENTATION PROOF OF CONCEPT

E.1. IMPLEMENTATION OVERVIEW

The TRAC M&S Behavior Validation Model (BVM) is an excellent example of a sound meta-data framework that was designed to capture information pertaining to the Modeling and Simulation Behavior key attributes, related relevant information and documentation. It also presents the associated conundrum of implementation, e.g. data management, document control, access, transparency, usability etc. This appendix describes the proof of concept implementation of the BVM as a Vector Relational Data Model (VRDM) in a Network Based Model Broker Architecture that is DoD Network Certified and under TRADOC ownership.

An implementation approach using VRDM was explored as a proof of concept for the BVM meta-model framework due to the following attributes:

- Low effort (cost).
- Configured (not programmed).
- Executable Model (solution and subsequent changes do not need recompiled).
- Network available models are configured through a web based configurable interface through any standard browser (e.g., Chrome, Safari, Firefox, Explorer).
- Transitionable to DoD network certified architecture that is a 'network model broker'. Dragon Pulse Information Network Architecture, DIACAP 2012 as a Type, MAC I Classified enclave. The Risk Management Framework in process for 2015.
- Extensible by SME (not programmers).
- Extends to COMBATXXI flat files and SQL backend, as well as any open application programming interface API.
- Allows data transparency and data driven navigation through the model.
- ERDC led research at TRAC and NPS:
- 2014 Thesis and Capstone awards for Cyber domain solutions.

- Demonstration of technology in multiple domains (ISR, gaming, KM, Cyber, Combat Search and Rescue and Joint Personnel Recovery (CSAR/JPR), healthcare, acquisition, Systems Engineering).
- Leads to an easily managed, dynamic and available (rules based) knowledge base capability with reduced man in the loop effort.

E.2. VRDM

The key independent concept objects are configured in VRDM as objects called "XTypes". The XTypes attributes are stored in virtual columns called "Elements". Each XType has a backend data source for storing or accessing "persistent" element data, and may also aggregate virtual elements from other XTypes. Each VRDM XType has a VRDM "Source" that is configured with permission and access to a network available data store. The backend data source in this case is simply MSSQL on the VRDM server. A one-to-one XType to MSSQL table construction allows for high-resolution access control to data (e.g. XType "MyVRDMData" may only consist of Elements ColumnA, ColumnG, ColumnZ of "YourDataSQLTable" that has ColumnA through ColumnZ) putting a hard stop on data flow or leakage from a VRDM model. A conceptual visualization is in Figure E.10. The loose-coupling (where system components have, or make use of, little or no knowledge of the definitions of other separate components) to the data source and back end store for both the model and the data in the model allows for a non-brittle (net easily extended due to many interdependent compiled systems) network distributed system. As depicted above, every Xtype and Xtype's Element has a related Source and Source Column that are related to the data through a related authorized Connection. This approach decouples the semantic layer form the syntactic layer and allows for System of systems (SoS) interoperability to be developed rapidly and easily. This is useful when owner of a database wants to share an explicit subset of tables with zero chance of spillage. In the case that there is total access allowed to a backend data source, the same high resolution controls may be implemented on data at the VRDM semantic layer, where domain vocabulary describes the model, but it is noted that it is now the responsibility of the VRMD data model(er) to specify access and usage of the data.

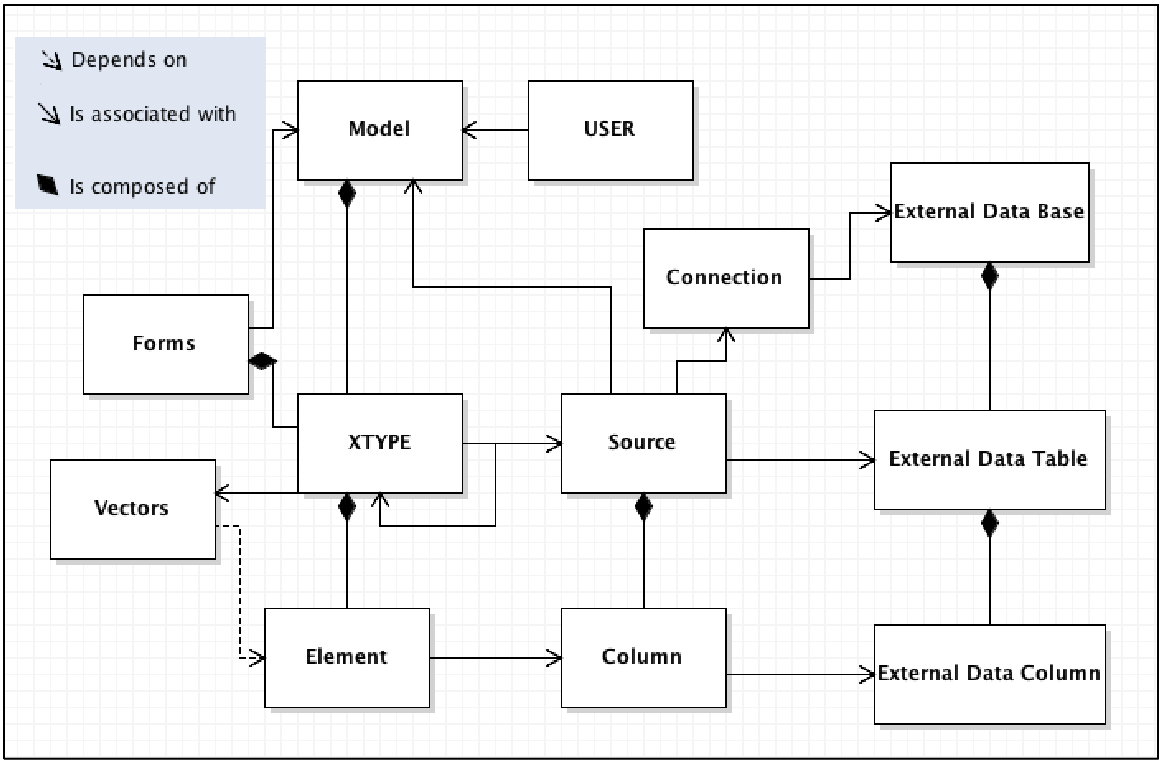


Figure E.10. Conceptualization of the GINA DPIMS Modeling Environment: *Conceptualization of the GINA DPIMS modeling environment. The core semantic model is defined by the XTypes and Vectors. The Vectors are also configured objects allowing for a configured executable model.*

Two other important concepts of VRDM are Vectors and Forms. Vectors are objects that specify relationships between XTypes. They may be considered a type of filter. Vectors are the innovative component to VRDM that allows for configuring executable models without the need for compiling. This is a subtle but important point, where as a SoS grows, a reconfiguration of the semantic model will work without the need for releasing new versions or restarting component systems.

VRDM Forms are the web form display of both the data and the data model that may be accessed in any standard web browser. Forms usually are constructed with an Authority Window, for navigation, a Resource Window for displaying object attributes, and a Collection Window, for displaying related object data. A search window is often added for large data sets. The forms may be configured to show all or none of the data from an XType, as well as allow change, add, and delete to the data or data model.

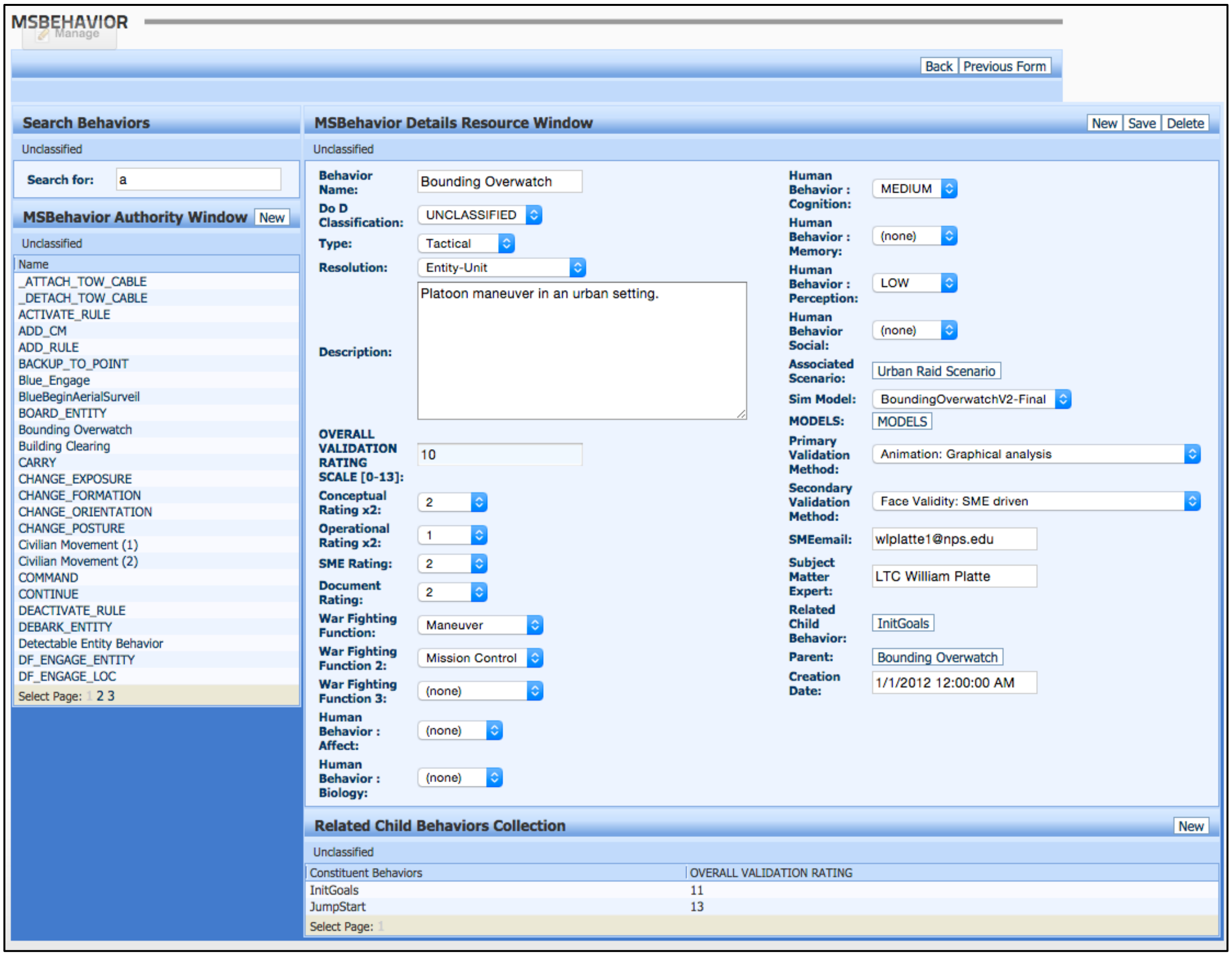


Figure E.11. VRDM Web Form: *A VRDM web form that has Search, Authority, Resource and Collection Windows for MSBehavior XType in the TRAC Behavior Validation Model. The Collection window, titled "Related Child Behaviors Collection" at the bottom shows data from a related XType object. This ability to configure and view unlimited data relationships offers navigable model, data, and transparency.*

E.2.1. VRDM Modeling Approach

The main aspect of creating a VRDM information model is to identify the key concepts, processes and outcomes of a desired of a system. In the case of the TRAC BV meta model, we started with the entire list and then distill the unique or key concepts that are to be understood as data objects unto themselves.

The steps for conceptual framework modeling the implementation of the TRAC Modeling and Simulation Behavior Validation meta model in VRDM:

1. Define Goal.

2. Problem Space Definition.
3. Functional Assessment: Determine problem space attributes.
4. Behavior Validation KM: Analytic Domain Model.
5. Refine and validate the information model.

Step 1: Define the Goal: Rate COMBATXXI behaviors and develop knowledge management capability for behavior validity with data transparency.

Step 2: The problem space definition in our case is the Behavior Validation and Knowledge Management.

Step 3: The Functional Assessment. The problem space attributes are determined and understood as key concepts. In our case, the TRAC meta-model that describes information pertinent and relevant to validation and pedigree of COMBATXXI behaviors is our starting point. The TRAC BVM attributes or meta-tags are listed below.

TRAC M&S BVMeta Model {
Behavior Name
Date
Behavior Description (to include trigger and action)
Parent Behavior
Component Behaviors
Associated Vignettes
Key Entities
Resolution
Behavior Type
Validation Technique
Warfighter Function
Human Behavior (Y/N)
Human Component (select and specify degree accounted for low/med/high)
Reference Doctrine
SME
Implementation Method
Project Study
Reference Model Implementation
Reference Model Version
Reference Documentation
POC
Security Distribution
Distribution
Conceptual Validation (0,2,4)
Operational Validation (0,2,4)
SME Developed (0-2)

Documented (0-3)
Total Score (0-13)
}

Step 4: We determine the VRDM key concepts from the key attributes of the TRAC meta-model. These are objects unto themselves that may be added to or changed in the future, and relate to the essential Behavior Validation object. For this proof of concept Behavior Validation VRDM implementation the following objects were determined and instantiated as the corresponding XTypes with related Sources and low coupled data sources shown in Figure E.12. Here, it is noted that there must be authorized access to the table, and that access to the table columns is specified column by column. This allows any subset of the table to be available to the VRDM model at the discretion of the data owner.

Behavior Validation Concept Object	GINA “XType” Object	GINA “Source” Object	Low Coupled Source (MSSQL Table)
Behavior	MSBehavior	MSBehavior	MSBehavior
Behavior Relationships	MSBehaviorRelationship	MSBehaviorRelationship	MSBehaviorRelationship
Model	MSBModel	MSBModel	MSBModel
DoD Classification	MSBDoDClassification	MSBClassification	MSBClassification
Behavior Type	MSBehaviorType	MSBehaviorType	MSBehaviorType
Documentation	MSBModelDocument	MSBModelDocument	MSBModelDocument
Rating 0-2	MSBTwoScale	MSBTwoScale	MSBTwoScale
Rating 0-3	MSBThreeScale	MSBThreeScale	MSBThreeScale
Rating High Med Low	MSBHiMedLow	MSBHiMedLow	MSBHiMedLow
Validation Technique	MSBValidation Technique	MSBValTechnique	MSBValTechnique
Vignette	MSBVignette	MSBVignette	MSVignette
Resolution	MSBResolution	MSBehaviorResolution	MSBehaviorResolution
CXXI ODB flat file FP_PROFILE_CT	MSBCXXI_JAN4_DFP_FS_P	MSBCXXIFS_Profile_CT	FS_PROFILE_CT *Imported from Apache ODB
CXXI ODB flat file OB_BEHAVIOR_CT	MSBCXXI_JAN4_DFP_behav	MSBCXXI_JAN4_DFP_behav	OB_BEHAVIOR_CT *Imported from Apache ODB
War fighting function	MSBWarFightingFunction	MSBWarfightingFunction	MSBWarfightingFunction

Figure E.12. Behavior Validation and Corresponding Xtypes, Sources and Back End Data: *Shows Behavior Validation concepts and corresponding Xtypes, Sources and back end data. The semantic simplicity allows rapid understanding of the relationships and components of the data model.*

Similar to the XType to Source to Low Coupled Source Table relationships shown in Figure 3, there are similar XType(Element) to Source(Column) to Low Coupled Source Table(column) relationships specified. Scripts speed the configuration of the VRDM model by propagating the naming conventions across the components.

A UML Class diagram of the Behavior Validation Model implementation is shown in Figure E.13. This illustrates the concept objects (XTypes) and their relationship (e.g. comprised of, related to) to other VRDM model objects.

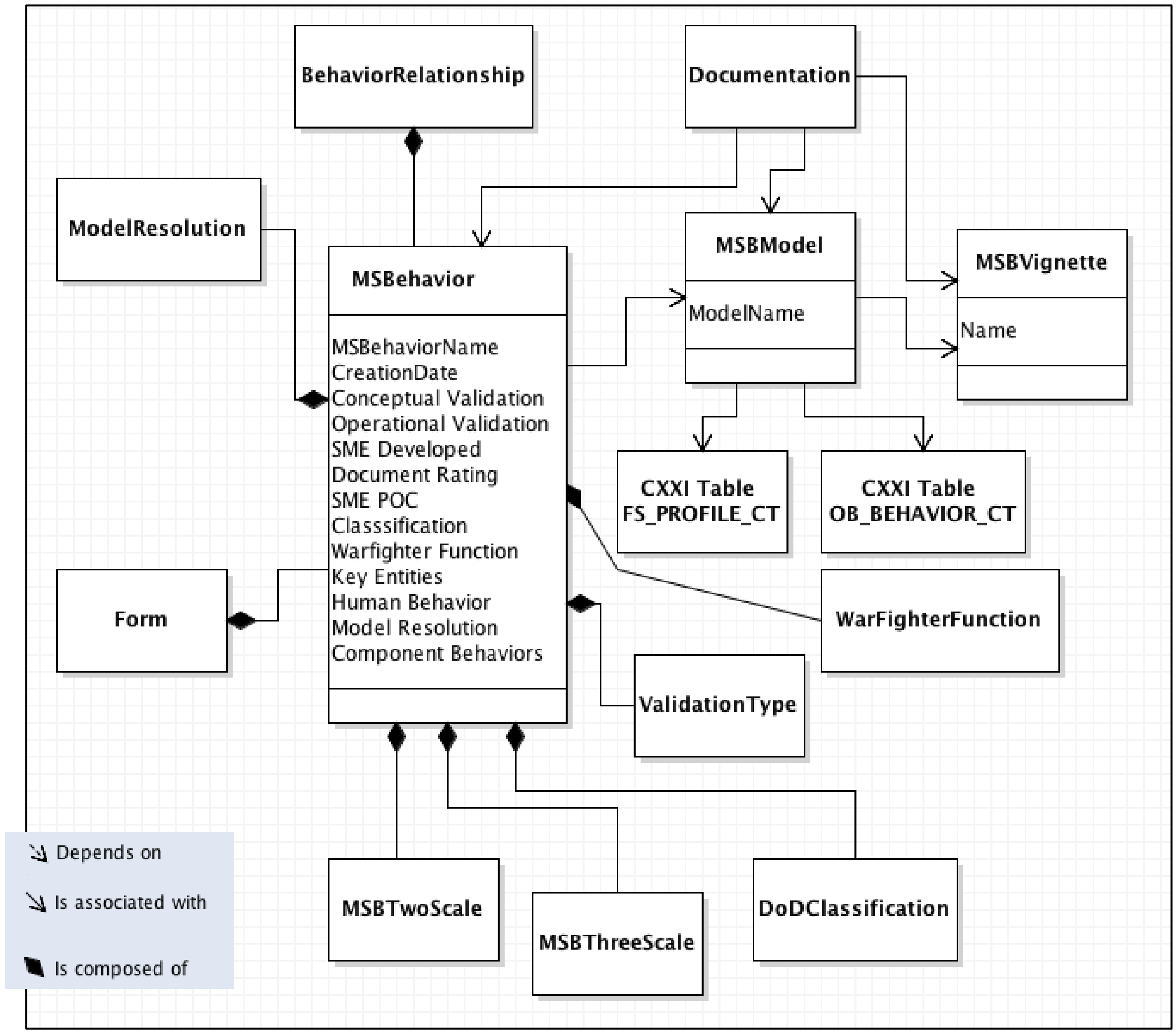


Figure E.13. Relationship Class Diagram*: A Class diagram that illustrates the relationships between key model component objects in the TRAC Behavior Validation Model. This perspective is from the MSBehavior XType object.*

Step 5: Refinement. This is self-explanatory, however it is noted that refinement is now just adding to tables and to objects, and specifying Vector Relationships. This allows Modeling and Simulation Analyst SME's to participate in fine-tuning the implementation of the VRDM without knowing about coding, source control etc.

E.2.2. VRDM Behavior Validation Model Overview

A quick overview of the TRAC MS Behavior Validation model implementation is included to describe some of the functionality. The search window (Figure E.14.) is configured to search on Behavior Name, SME, or description. Figures E.6 through E.16 describe the interaction with the data model. This was a limited proof of concept that did not exploit any of the Services capabilities inherent to the VRDM framework, which allow for process and system behavior execution.

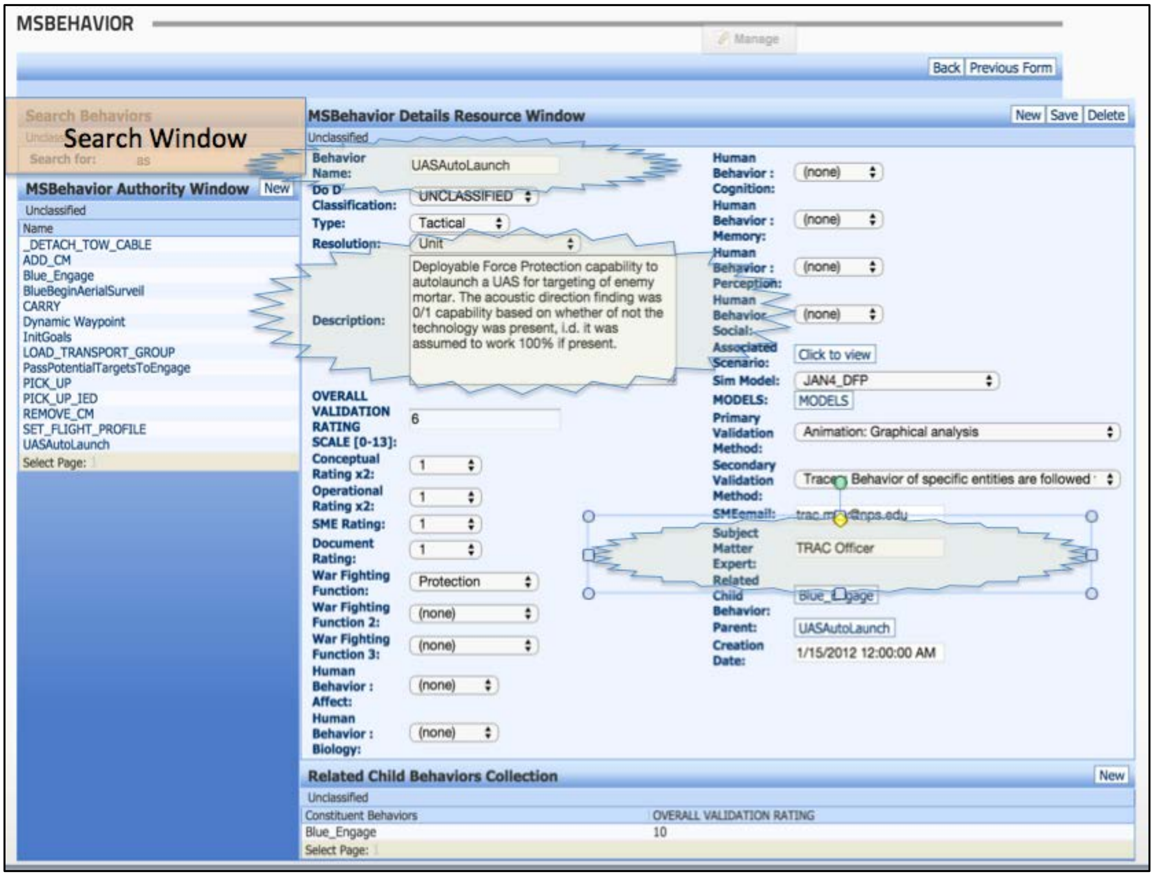


Figure E.14. Web Form Search Capacity: *The search window is shown in the top left of the MSBehavior Form. It is configured to search on Behavior Name, Description Field, or Subject Matter Expert.*

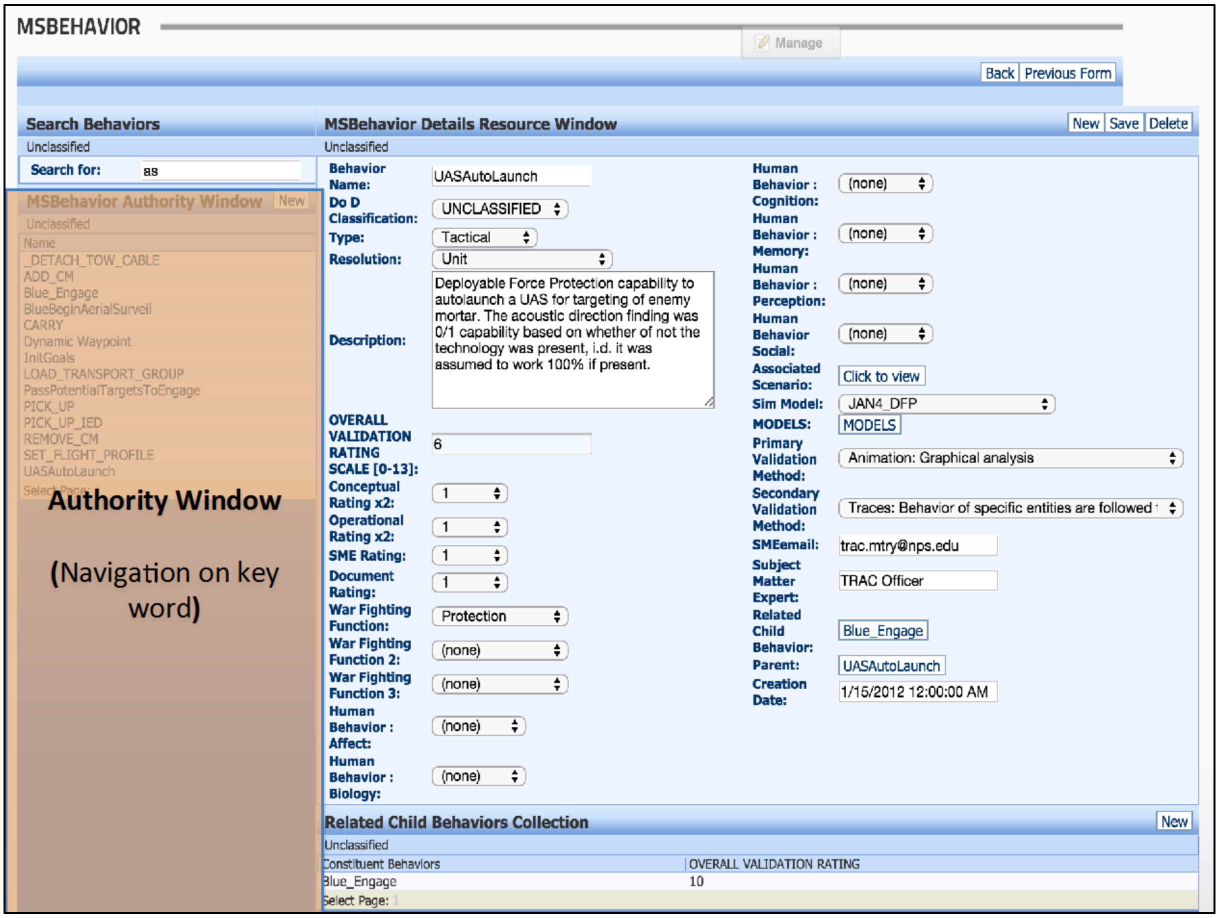


Figure E.15. Web Form's Authority Window: *A Form's Authority window is usually configured to be on the left hand side of the Form and displays the results from the search and allows selection of the key object, that is hyper linked to the Resource's metadata display window on the upper right.*

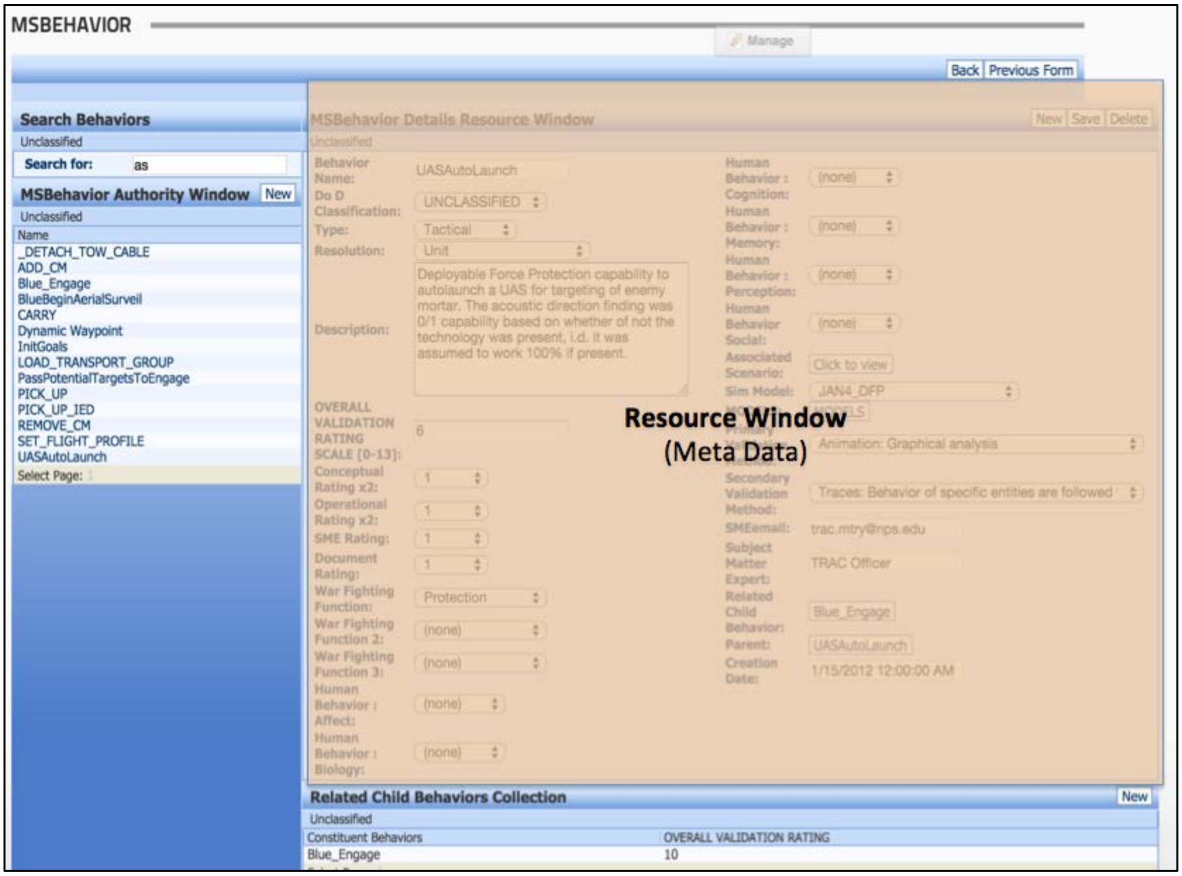


Figure E.16. Web Form's Resource Window: *The Form's Resource Window, displays and allows management of the data object's related metadata.*

Figure E.17. Web Form's Collection Window: *The Collection window displays data related to the data object or filtered on data object attribute.*

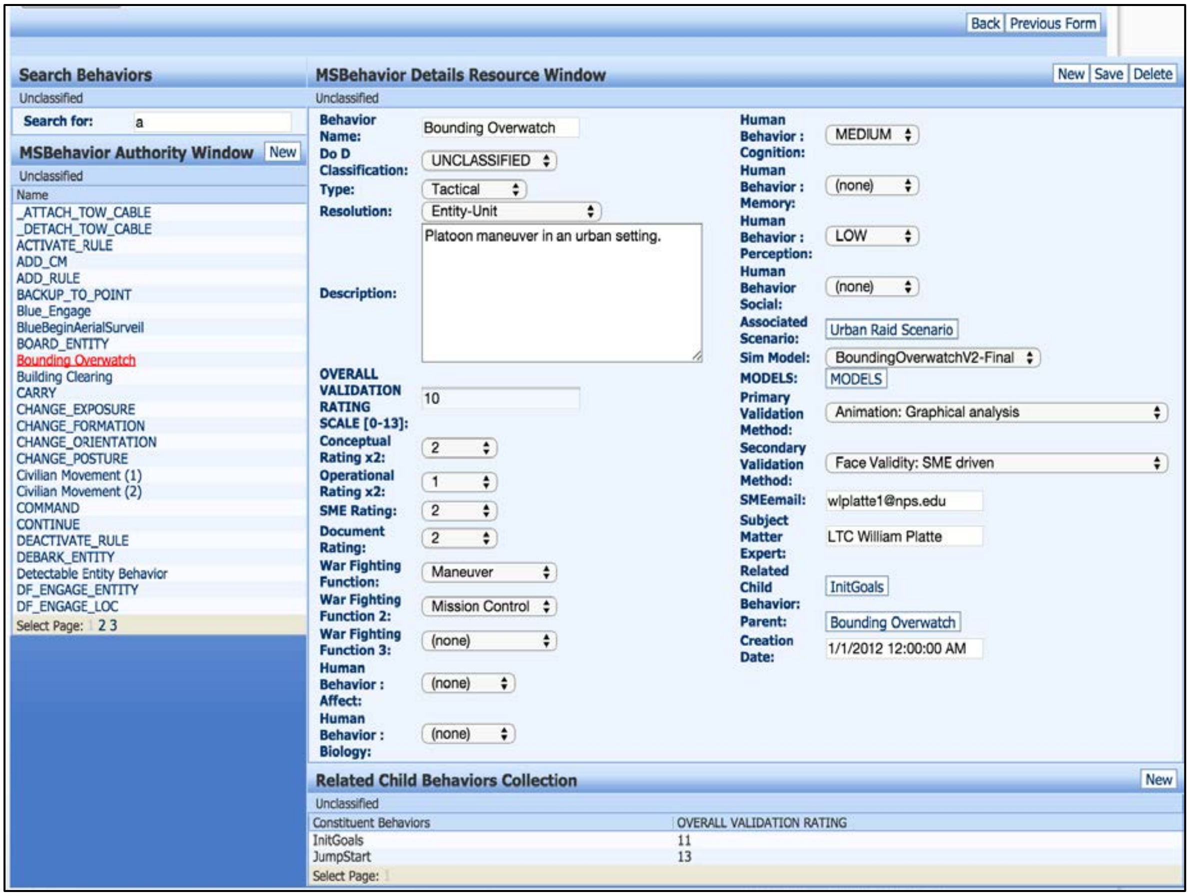


Figure E.18. Search Results and Bounding Overwatch Behavior Meta Data*:* *The MSBehavior Form above shows results from searching on "a". A long list of matches is shown in the Authority Window on the left side. The "Bounding Overwatch" behavior is selected, and the meta data is displayed for it in the Resource Window. The Collection Window shows the related Child Behaviors. The data attributes may be changed via drop down lists and text fields. New Behaviors may also be created with the "New" button and the rectangular blue buttons navigate to related XTypes.*

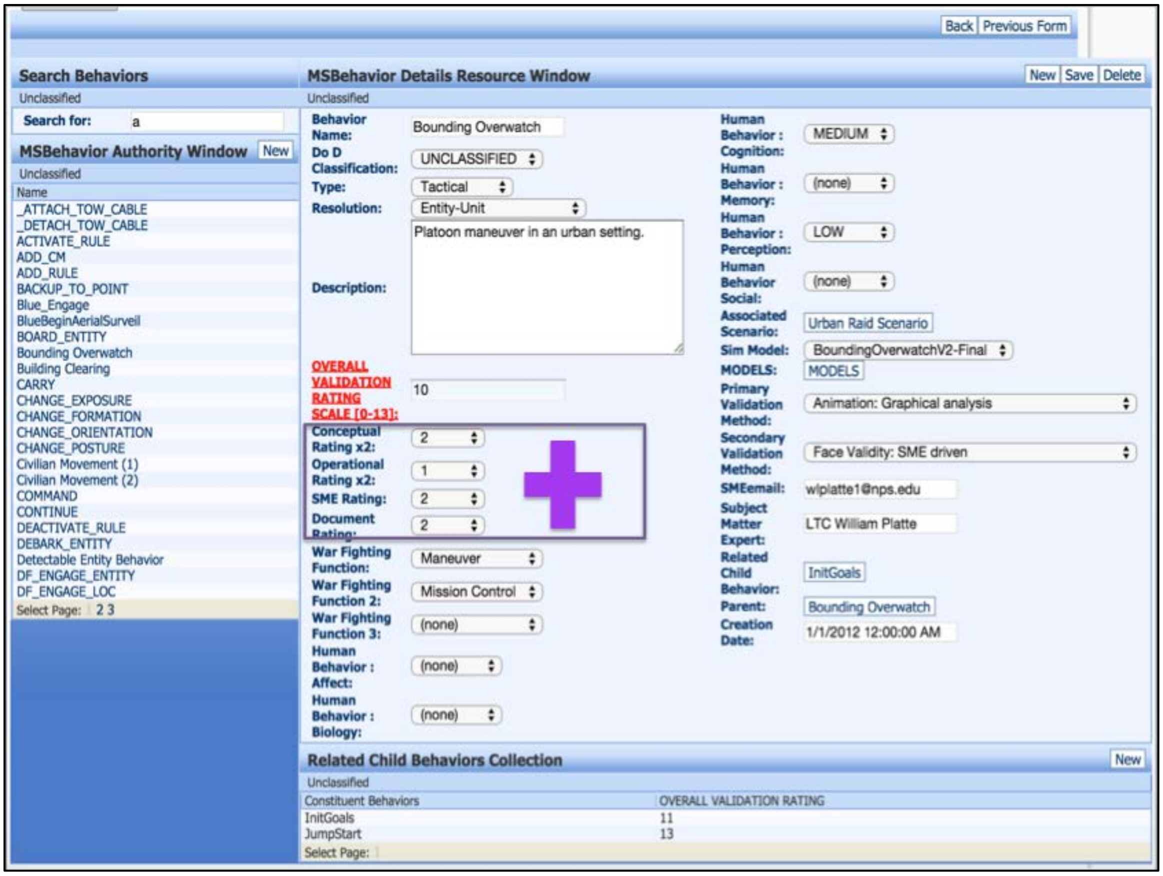


Figure E.19. Behavior Validation Rating Scale Web Visualization: *The individual rating categories are specified by drop down windows and the sum of the values is dynamically represented in the "Overall Validation Rating Scale (0-13)".*

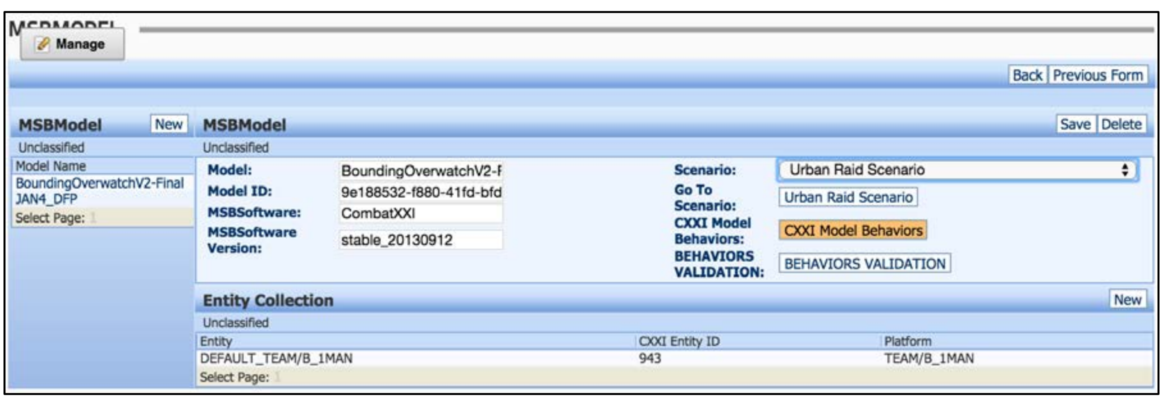


Figure E.20. MODELS Button Navigation Window: *The "MODELS" button for the MSBehavior Form navigates to the Form for the associated MSBModel. Here new modes may be added. From this form, the CXXI Model Behaviors button (highlighted) navigates to the related XType representing the CXXI behavior table data.*

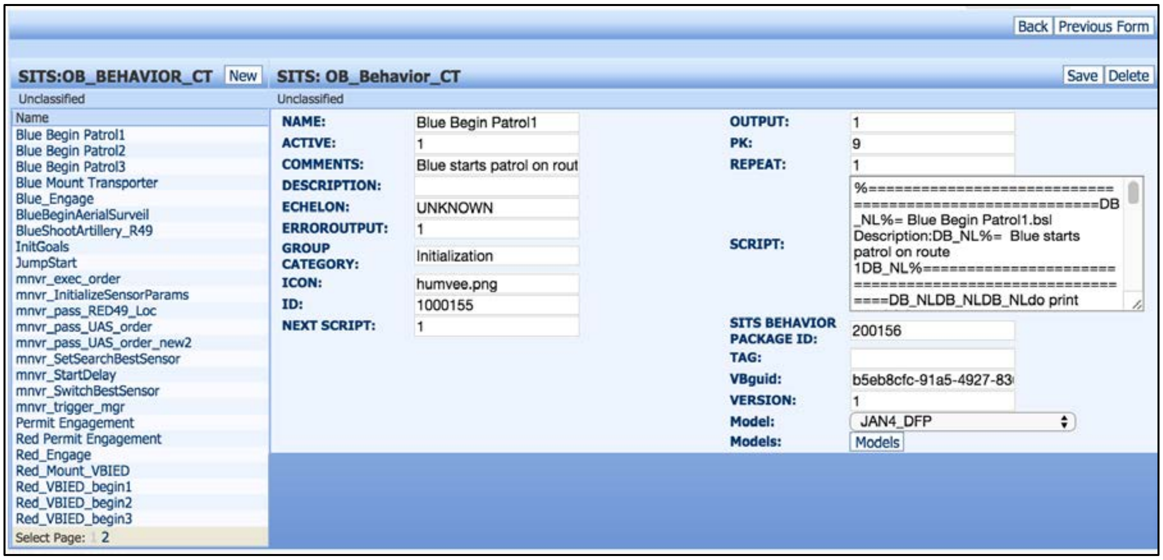


Figure E.21. Related Behaviors and Data Tables: *The Form for the COMBATXXI OB_BEHAVIOR_CT shows all of the related model behaviors and COMBATXXI data table content. The "Models" navigate button allows navigation back to the Behavior Model via the Model Form.*

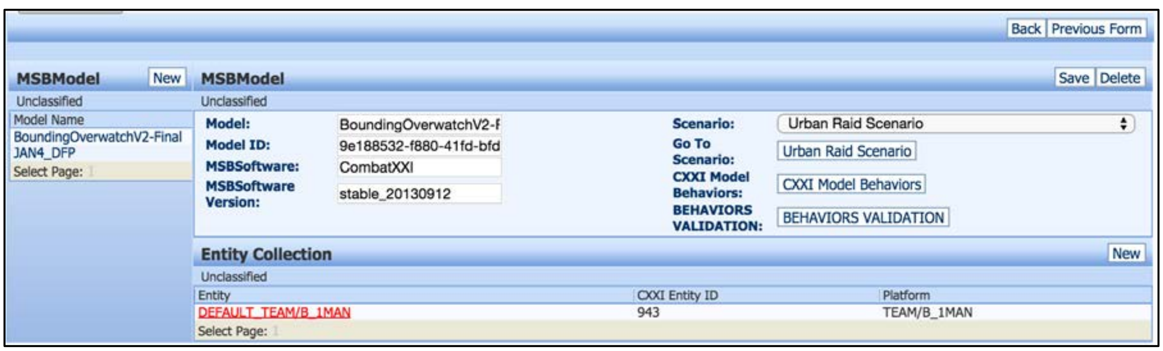
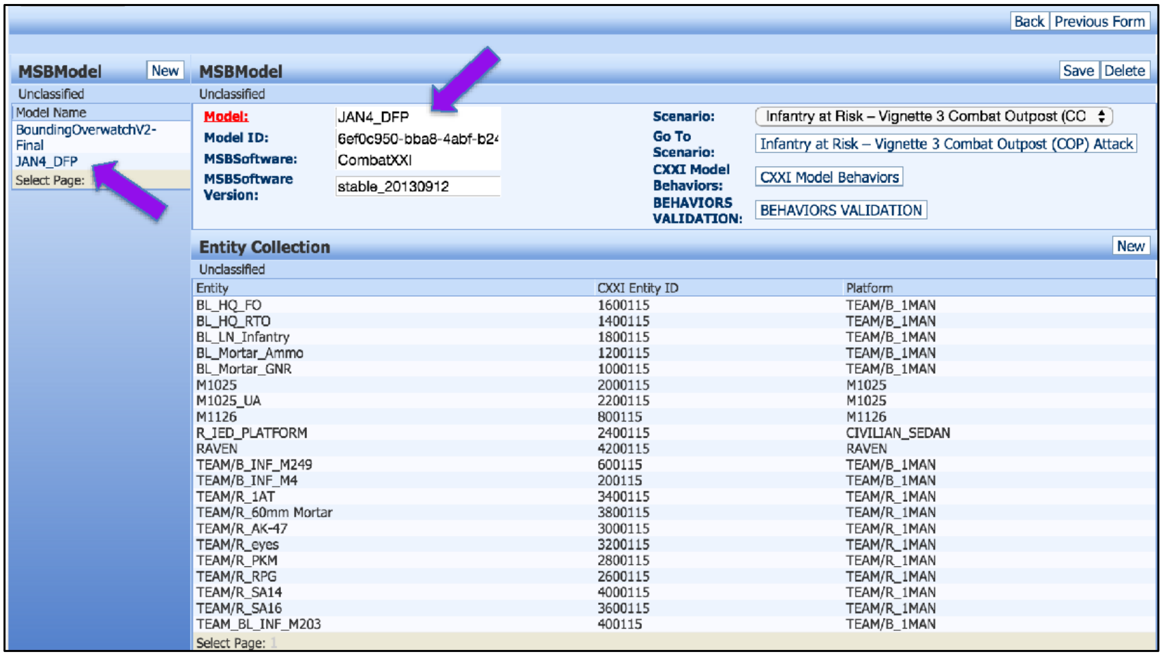


Figure E.22. Models Form and Relationship to Bounding OverwatchV2: *On the Models Form, the Collection window shows Entities that are used in the model. This information is collected from another COMBATXXI ODB flat file that was ingested. Here, the collection window is configured to show Entity Name, Code and Platform associated with the model (Bounding OverwatchV2).*



Entity	CXXI Entity ID	Platform
BL_HQ_FO	1600115	TEAM/B_1MAN
BL_HQ_RTO	1400115	TEAM/B_1MAN
BL_LN_Infantry	1800115	TEAM/B_1MAN
BL_Mortar_Ammo	1200115	TEAM/B_1MAN
BL_Mortar_GNR	1000115	TEAM/B_1MAN
M1025	2000115	M1025
M1025_UA	2200115	M1025
M1126	800115	M1126
R_IED_PLATFORM	2400115	CIVILIAN_SEDAN
RAVEN	4200115	RAVEN
TEAM/B_INF_M249	600115	TEAM/B_1MAN
TEAM/B_INF_M4	200115	TEAM/B_1MAN
TEAM/R_1AT	3400115	TEAM/R_1MAN
TEAM/R_60mm Mortar	3800115	TEAM/R_1MAN
TEAM/R_AK-47	3000115	TEAM/R_1MAN
TEAM/R_eyes	3200115	TEAM/R_1MAN
TEAM/R_PKM	2800115	TEAM/R_1MAN
TEAM/R_RPG	2600115	TEAM/R_1MAN
TEAM/R_SA14	4000115	TEAM/R_1MAN
TEAM/R_SA16	3600115	TEAM/R_1MAN
TEAM_BL_INF_M203	400115	TEAM/B_1MAN

Figure E.23. Collection of Entities from COMBATXXI Apache ODB File: *The selection of the JAN4_DFP model from the Authority window yields the model meta data and the Collection of entities from COMBATXXI Apache ODB file.*

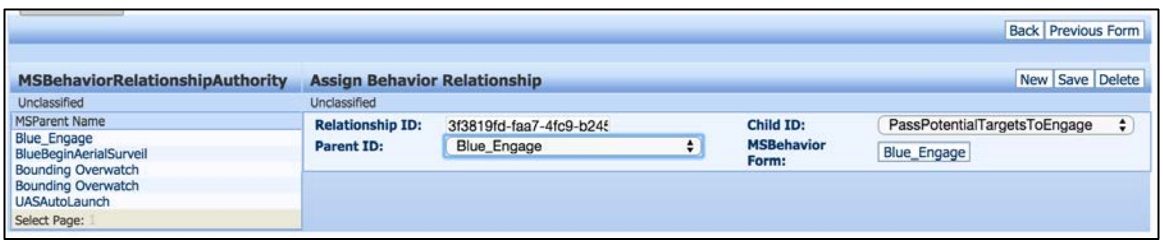


Figure E.24. Related Child Behaviors Creation: *Selecting the "Related Child Behaviors" from the MSBehavior form navigates to the "Assign Behavior Relationship" Form. Here new parent child relationships are specified from drop down lists of the available behaviors.*

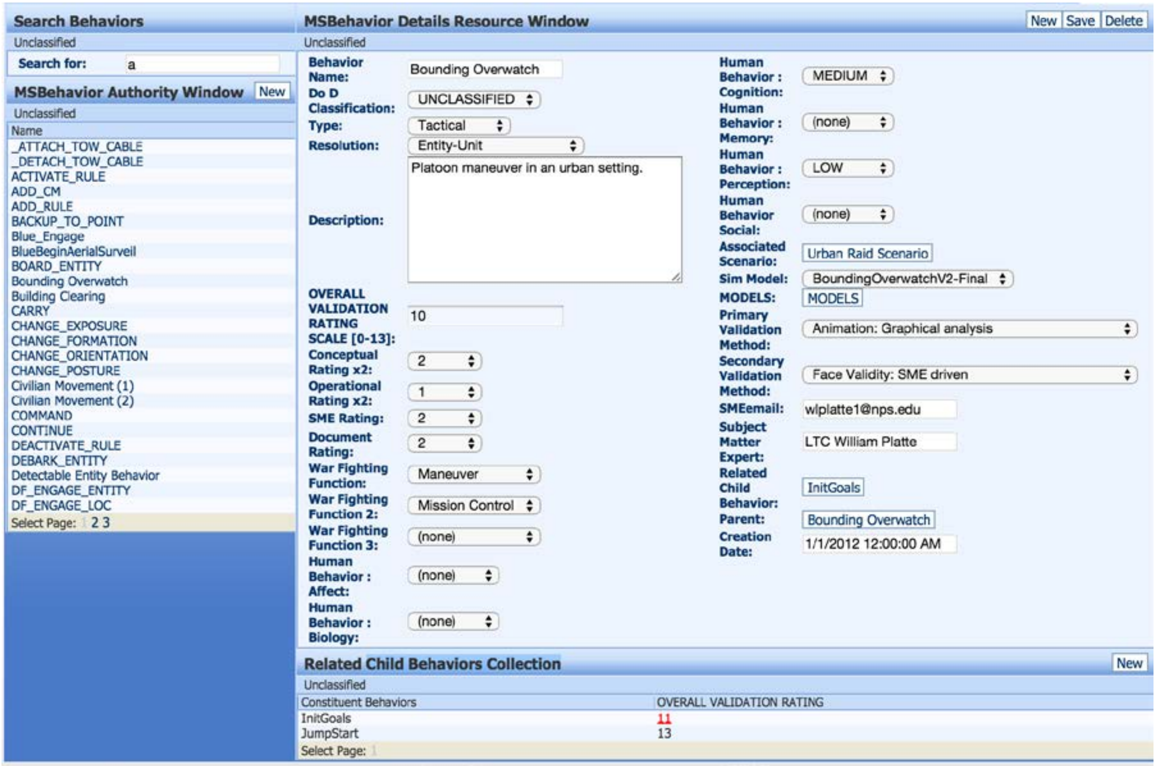


Figure E.25. Child Behaviors with Associated Ratings: *The MSBehavior Form shows the related child behaviors along with the associated ratings in the Collection window.*

LIST OF REFERENCES

ADRP 3-0. (2012, MAY). *UNIFIED LAND OPERATIONS*. Department of the Army.

Alt, J. K. (2008). *A Reference Model of Soldier Attention and Behavior*. Monterey: TRAC MTRY & Naval Postgraduate School.

AR 5-11.30. (2005). *Management of Army Models and Simulations*. Department of the Army.

Birta, L. G., & Ozmizrak, N. F. (1996, January). A Knowledge-Based Approach for the Validation fo Simulation Models: The Foundation. *ACM Transactions on Modeling and Computer Simulation*, 76-98.

BOCME Scenario Builders Guide. (2011). *Basic Observer, Communication, Move and Engage Scenario*. TRAC WSMR - Department of the Army.

Flournoy, R. D. (2002, April). Leveraging Human Behavior Modeling Technologies to Strengthen Simulation - Based C2 System Acquisition. Bedford, MA: MITRE Corporation.

FM 3-21.8. (2007, MARCH). *Infantry Rifle PLT and SQD*. Department of the Army.

Goerger, S. R., McGinnis, M. L., & Darken, R. P. (2005). *A Validation Methodology for Human Behavior Representation Models*. West Point: Operations Research Center of Excellence.

Goerger, S. R., McGinnis, M. L., & Darken, R. P. (2005). A Validation Methodology for Human Behavior Representation Models. *JDMS Volume 2, Issue 1*, 5-17.

Johns Hopkins University Applied Physics Laboratory. (2011, May 18). *M&S VV&A RPG Core Document: Introduction*. Retrieved August 2014, from VV&A Recommended Practices Guide : https://msenterprise.jhuapl.edu/RPG/

Law, A. (2007). *Simulation Modeling and Analysis. 4th ed.* New York: McGraw-Hill.

Sargent, R. G. (2011). Verification and Validation of Simulation Models. *Winter Simulation Conference*, 183-198.

Silverman, B. G., Bharathy, G., O'Brien, K., & Cornwell, J. (2006). Human Behavior Models for Agens in Simulators and Garmes: PartII Gamebot Engineering with PMFserv. *Presence: Teleoperators and Virtual Environments Volume 15*, 163-185.

TRAC WSMR. (2012, September 13). COMBATXXI Behaviors Overview. White Sands Missile Range, NM.

TRAC WSMR. (2014). *COMBATXXI Users Guide: Behavior Rating System*. Retrieved August 2014, from CONFLUENCE: https://cxxi.wsmr.army.mil/confluence

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