NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

A METHODOLOGY TO ASSESS URBANSIM SCENARIOS

by

Brian D. Vogt

September 2012

Thesis Advisor:	Joseph Sullivan
Second Reader:	Jonathan Alt

This thesis was performed at the MOVES Institute
Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704–0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE A Methodology to Assess UrbanSim Scenarios			**5. FUNDING NUMBERS**
6. AUTHOR(S) Brian D. Vogt			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943–5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A______.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

Turn-based strategy games and simulations are vital tools for military education, training, and readiness. In an era of increasingly constrained resources and expanding demand for training solutions, the need for validated, effective solutions will increase. Appropriate performance feedback is an important component of any training solution. Current methods for designing and testing the performance feedback provided in turn-based simulation are limited to well-structured problems and do not adequately address ill-structured problems that better replicate problems facing military leaders in today's complex operating environment. This thesis develops and explores new methods for assessing the feedback mechanisms of turn-based strategy games. Using UrbanSim, a game for training strategic approaches to COIN operations as an exemplar, this thesis developed and explored two unique methods for evaluating the reward structure of the UrbanSim scenarios. The first method evaluates different student strategies using a batch-run method. The second method uses a reinforcement-learning algorithm to explore the decision space. These scenario evaluation methodologies are shown to be able to provide insights about a game's performance feedback mechanism that was not previously available. These methodologies can be used for formative evaluation during game scenario development. Additionally, these evaluation methodologies are generalizable to other training and education games that focus on ill-structured problems and decision-making at discrete intervals.

14. SUBJECT TERMS UrbanSim, Games for Training, Reinforcement-learning, Performance Feedback Assessment			**15. NUMBER OF PAGES** 101
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01–280–5500

Standard Form 298 (Rev. 8–98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

A METHODOLOGY TO ASSESS URBANSIM SCENARIOS

Brian D. Vogt
Major, United States Army
B.S., University of Wisconsin–Oshkosh, 1996

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN
MODELING, VIRTUAL ENVIRONMENTS, AND SIMULATION (MOVES)**

from the

**NAVAL POSTGRADUATE SCHOOL
September 2012**

Author: Brian D. Vogt

Approved by: Joseph Sullivan
Thesis Advisor

Jonathan Alt
Second Reader

Christian Darken
Chair, MOVES Academic Committee

Peter J. Denning
Chair, Computer Science Academic Committee

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Turn-based strategy games and simulations are vital tools for military education, training, and readiness. In an era of increasingly constrained resources and expanding demand for training solutions, the need for validated, effective solutions will increase. Appropriate performance feedback is an important component of any training solution. Current methods for designing and testing the performance feedback provided in turn-based simulation are limited to well-structured problems and do not adequately address ill-structured problems that better replicate problems facing military leaders in today's complex operating environment. This thesis develops and explores new methods for assessing the feedback mechanisms of turn-based strategy games. Using UrbanSim, a game for training strategic approaches to COIN operations as an exemplar, this thesis developed and explored two unique methods for evaluating the reward structure of the UrbanSim scenarios. The first method evaluates different student strategies using a batch-run method. The second method uses a reinforcement-learning algorithm to explore the decision space. These scenario evaluation methodologies are shown to be able to provide insights about a game's performance feedback mechanism that was not previously available. These methodologies can be used for formative evaluation during game scenario development. Additionally, these evaluation methodologies are generalizable to other training and education games that focus on ill-structured problems and decision-making at discrete intervals.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ALC 2015 – Army Learning Concept 2015

ALM 2015 – Army Learning Model 2015

BC2010 – Battle Command 2010 Intelligent Tutoring System

CoE – Center of Excellence

ComMentor – Command Mentoring Intelligent Tutoring System

DARPA – Defense Advanced Research Projects Agency

DP – Deliberate Practice

DQ-C – Direct-Q Computation

ELM – Experiential Learning Model

FM - Field Manual

IED – Improvised Explosive Device

ILE – Intermediate Level Education

LOE – Line of Effort

MMOG – Massively Multiplayer Online Game

MC3 – Maneuver Captain's Career Course

MSCCC – Maneuver Support Captain's Career Course

PME – Professional Military Education

POMDP – Partially Observable Markov Decision Problem

PsychSim – Psychological Simulation

RDECOM – U.S. Army Research Development Engineering Command

S2 – Intelligence Officer

S3 – Operations Officer

SCP – School for Command Preparation

STTC – Simulation Training Technology Center

TAO – Tactical Action Officer

TAO ITS – Tactical Action Officer Intelligent Tutoring System

ACKNOWLEDGMENTS

First, I need to thank God for the opportunity to study at the Naval Postgraduate School and work with such great people. Second, I need to thank my wife, who has been a great teammate and cheerleader, and my daughters, who not only endure the Army lifestyle but thrive in it. Third, I need to thank CDR Joseph Sullivan and LTC Jon Alt, who kept me motivated, focused, and grounded through many impromptu office discussions and late nights in the TRAC-Monterey Combat Models Lab.

The next group of folks were tremendously helpful in various ways. MAJ Shane Price, my battle-buddy through many challenging courses, is a great friend and sounding board for ideas. LTC Glenn Hodges helped me shape the direction of this thesis. Curt Blais encouraged me to expand a class project into this thesis and facilitated an opportunity to present this topic at a conference that proved to be very fruitful. The MOVES Institute instructors were genuinely interested in my education and provide a world-class education. Steve Hebert helped my get my head wrapped around Python. Dr. Bob Pokorny, graciously provided insight and experience concerning UrbanSim scoring and other facets of using games for training and education. TRAC-Monterey, especially Sandra Lackey and Jimmy Liberato, were very supportive by allowing me to use a cubicle in the combat models lab for several months. Tim Wansbury, RDECOM, was gracious by providing unvarnished insight about the development of UrbanSim. Sowmya Ramachandran and Jim Ong, of Stottler-Henke Associates, provided great insight about successes and challenges with developing and understanding intelligent tutoring systems.

People I have not met face to face also invaluably assisted this effort. Stacy Marsella and David Pynadath, from the USC ICT, created PsychSim and provided a tremendous amount of assistance and insight through countless e-mails. I do not think my study would have been possible if it were not for their code and assistance.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

> The Army requires the capability to develop adaptive digitized learning products that employ artificial intelligence and/or digital tutors to tailor learning to the individual Soldiers'' experience and knowledge-level and provide a relevant and rigorous, yet consistent, learning outcome. (U.S. Army, 2011)

The use of games and gaming to educate is certainly not new. Games have been used in educational settings for many years with varying levels of success. Many times these games have focused on well-defined problems such as math, science, and procedural trainers. The reward structure of these types of games can be directly validated if they reward the student with the one correct answer or solution. However, there has been an increased desire to use games to train and educate students to perform well in ill-defined problem areas. Ill-defined problems are characterized as having more than one correct, or acceptable, solution. Validation of games that address ill-defined problems is inherently more difficult than well-defined problems. One of the challenges in the application of complex agent based games built for training and education is the verification that the intended learning outcomes are being reinforced by the training system, and likewise that undesired behaviors are not being rewarded. This thesis will address this challenge with two methods. The first method is a batch run method that bins actions into different strategies and each strategy is tested numerous times. The second method uses a reinforcement-learning agent that explores different strategies and provides feedback about how the strategies are rewarded.

The U.S. Army's use of a game called UrbanSim provides an example of such a use case. UrbanSim is a turn-based strategy game that is designed to train leaders in executing battle command in complex environments focused on counterinsurgency and stability operations (Wansbury, Hart, Gordon, & Wilkinson, 2010). UrbanSim was developed and fielded by the U.S. Army as a tool to support educational objectives concerning counterinsurgency operations at the School of Command Preparation at Fort Leavenworth, Kansas. The front-end analysis of UrbanSim and the associated scenarios used for training were based on extensive interviews with battalion and brigade commanders that returned from Iraq (Wansbury, Hart, Gordon, & Wilkinson, 2010). After collecting and collating this information, the development team presented it to the Combined Arms Center at Ft. Leavenworth to ensure the principles were in line with doctrine and current counterinsurgency principles. Next, the development team produced UrbanSim, with PsychSim as the underlying simulation. UrbanSim testing primarily focused on software stability to ensure it was able to operate on the intended hardware platforms. A reasonable method to evaluate the scenarios and the performance feedback mechanisms was not readily available to the development team (Wansbury, 2011).

There is limited direct evidence to support that the scenarios developed and fielded supported the educational objectives. That is to say, that the embedded performance feedback mechanisms within UrbanSim has not been evaluated to ensure students were guided through rewards and penalties to achieving better understanding of COIN operations. The development team assumed risk in this area because UrbanSim was intended to be used in the classroom with an instructor. If the results of actions in the game did not seem correct, or falsely rewarded poor decisions, the instructor was able to give verbal feedback to overcome this apparent shortcoming of the UrbanSim scenario performance feedback. Additionally, scenario validation did not seem feasible at the time of fielding due to the vast number of possible ways to play the game. The use of UrbanSim has grown from a simulation to support Fort Leavenworth's School of Command Preparation under the supervision of an experienced instructor to being used at Captain Career Courses, Non-Commissioned Officer Academies, Service Academies, as well as available to all Soldiers via the Army Military Gaming website. These expanded uses reduce the role of an experienced instructor that can guide students when the results of the game are contrary to desired learning objectives. Therefore, it is situations like this that it is becoming increasingly important to ensure the performance feedback mechanisms in training and educational games properly reward good performance and penalize poor student performance.

UrbanSim is a good test-case of a larger problem with simulations and games for education. UrbanSim was designed for use with an instructor guiding the learning experience. However, UrbanSim is now fielded and available without instructors. If we can figure out what is missing or needed to effectively use UrbanSim without instructors, we will make progress toward designing effective simulations without instructors.

A. RESEARCH QUESTIONS

This thesis will address the overarching research question:

Can batch-running or using a reinforcement-learning approach provide useful insights about the performance feedback mechanism of UrbanSim?

Within the overarching research effort, this thesis will address the following research questions:

- Does UrbanSim’s performance feedback system support the stated learning objectives?
- Does the scenario reward a ‘‘Clear, Hold, Build’’ strategy better than the other strategies?
- Does the scenario reward student actions that are exclusively legal over student actions that are a mixture of legal and illegal actions?
- Does the scenario reward student actions that are a mixture of lethal and non-lethal actions over exclusively lethal or exclusively non-lethal?
- Is the performance feedback provided to the learner strong enough to differentiate between optimal and non-optimal strategies?

B. BENEFITS OF THIS STUDY

The two primary benefits of this study are 1) provide an analysis of a currently fielded UrbanSim scenario and 2) inform a generalizable method to analyze games that seek to educate and train students about ill-defined problems.

The UrbanSim scenarios used across the Army today have not been explicitly validated to ensure that good actions are rewarded and poor actions are penalized in the performance feedback mechanisms. This study seeks to address this identified shortfall.

There is great potential for game and simulation development to address the wider field of ill-defined problems and provide very efficient means to train and educate leaders concerning complex environments. However, validation of these types of games and simulations can be rather daunting. This study intends to address this challenge with a generalizable approach to validate games and simulations that seek to train and educate about ill-defined problems.

This study fully supports the vision outlined in the Army Learning Concept 2015 by providing a method to evaluate UrbanSim scenarios as they relate to the specified training and educational objectives. Additionally, this study provides a generalizable approach to validate training and educational game scenarios for a specific class of ill-defined problems.

C. THESIS ORGANIZATION AND TABLE OF CONTENTS

- Chapter I: Introduction. This chapter describes the problem, lists the research questions, and defines the scope and benefits of this study.
- Chapter II: Background. This chapter provides a literature review for the study. This review includes current literature on doctrine, experiential learning model, deliberate practice, performance

feedback, game based training, current intelligent tutoring systems, and a description of UrbanSim and PsychSim

- Chapter III: Methodology. This chapter describes how the research team designed the experiments.
- Chapter IV: Results and Discussion. This chapter contains the results of the experiments and an interpretation of those results.
- Chapter V: Recommendations. This chapter provides an overall assessment, methods to evaluate other scenarios, limitations of this methodology, and recommends future work for assessing scenarios to train ill-defined problem solving.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

A. CHANGES IN CURRENT OPERATING ENVIRONMENT THAT NECESSITATE CHANGES IN THE TRAINING AND EDUCATION ENVIRONMENT

Since 2001, the U.S. military has been primarily involved in counterinsurgency and stability operations as opposed to the traditional major combat operations that dominated training and education within the military for the preceding two decades. Major combat operations are characterized by overwhelming combat power applied at decisive points on the battlefield to impose the commander’s will and change the environment to the desired end state (U.S. Army, 2011). Conversely, counterinsurgency and stability operations are characterized by carefully planned and executed combat and stability operations used to facilitate the main effort of supporting the population (U.S. Army, 2006). While major combat operations create an immediate change to an environment, counterinsurgency and stability operations creates a lasting, sustainable solution that is satisfactory to our goals and objectives.

The UrbanSim training package was developed in direct response to the unique challenges of counterinsurgency and stability operations. Senior leaders within the Army identified educational and training shortcomings of Army leaders to effectively operate in such a complex and challenging environment. To be successful, leaders could not simply fight their way to success, but rather use a wide range of operations to help set the conditions for the host nation population to develop their police and military forces, government agencies, and social order.

B. ARMY FIELD MANUAL, FM 7-0, *TRAINING UNITS AND DEVELOPING LEADERS FOR FULL SPECTRUM OPERATIONS*

FM 7-0 is the Army’s capstone document on training and educating the Army to meet the challenges of the contemporary operating environment. FM 7-0 provides specific guidance about training and educating leaders. First, it is recognized that “time is the scarcest resource when we confront training” (U.S. Army, 2011). Therefore, when applying Ericsson’s principles of deliberate practice, the Army must seek, develop, and implement methods of training and education that efficiently use the scarce resource of time. Second, “Among the three aspects of leader development—training, education, and experience—experience is the most direct and powerful. Subordinates learn by doing. Lessons learned while making mistakes can be the best way to improve as a leader” (U.S. Army, 2011) This direct observation about experiential learning also implies that leaders must learn from the consequences of their actions and that making mistakes can be an effective tool to train and educate. Third, the Army training management cycle of plan, prepare, execute, while always assessing and providing feedback, is similar to Kolb’s experiential learning model of 1) a concrete experience, 2) reflective observation, 3) abstract conceptualization, and 4) active experimentation.

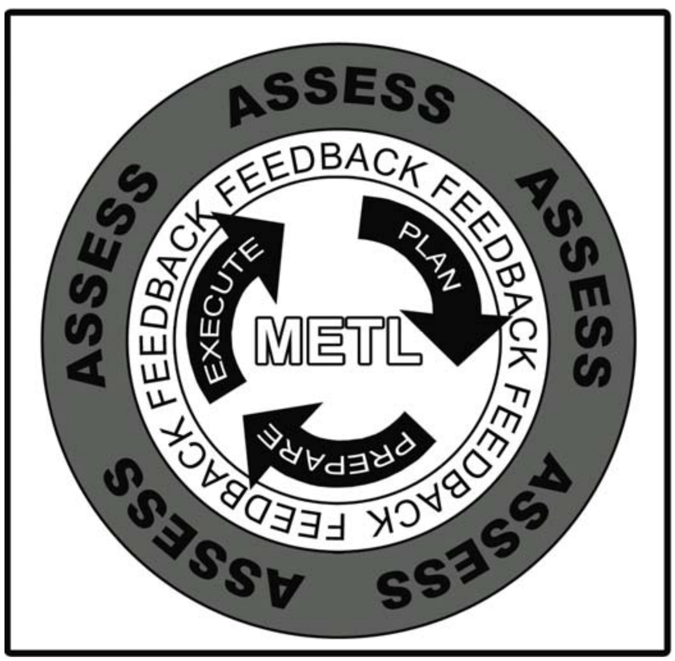

Figure 1. The Army training management model (From U.S. Army, 2011)

Last, FM 7-0 describes the three domains of training and education. They are institutional, operational, and self-development domains. There is considerable simulation support for institutional and operational domains development but few simulation tools to assist with individual professional development. Recent efforts, as outlined in the Army Learning Model 2015 seek to address this identified shortcoming.

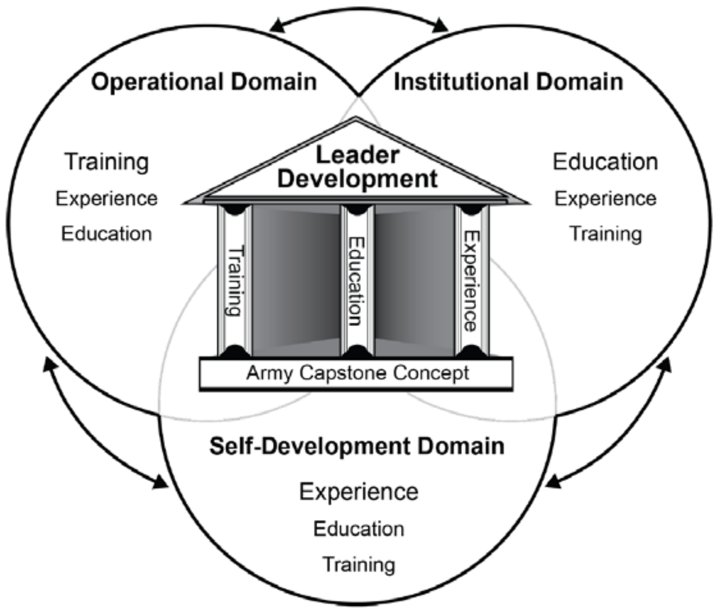

Figure 2. The Army's leader development model (From U.S. Army, 2011)

C. ARMY LEARNING MODEL 2015

The Army Learning Model 2015 (ALM 2015) is described in TRADOC Pamphlet 525-8-2, *The Army Learning Concept for 2015* (ALC 2015). ALM 2015 "seeks to improve our learning model by leveraging technology without sacrificing standards so we can provide credible, rigorous, and relevant training and education for our force of combat seasoned Soldiers and leaders" (U.S. Army, 2011). ALC 2015 describes the current learning environment with the Army learning institutions as:

> based on individual tasks, conditions, and standards, which worked well when the Army had a well-defined mission with a well-defined enemy… Mandatory subjects overcrowd programs of instruction (POIs) and leave little time for reflection or repetition needed to master fundamentals. Passive, lecture-based instruction does not engage learners or capitalize on prior experience. (U.S. Army, 2011)

ALM 2015 describes that the Army desires to shift to addressing the inherently ill-defined problems that our Army currently faces and will increasingly face in the future. Additionally, it calls for a capability for Soldiers to reflect on their learning and be able to repeat the exercises to master fundamentals. The ALC 2015 recognizes that rote memorization used in the past no longer meets the needs of the Army. These concepts are aligned with current learning theories and practice. Specifically, they reflect the ideas of Ericsson et al.'s (1993) deliberate practice and Clark's (2008) description how to develop and maintain expertise.

The ALC 2015 describes characteristics of its leaders as adaptable, able to operate in decentralized operations, and masters of the fundamentals. These characteristics are not natural abilities, but rather developed through education, training, and most importantly through deliberate practice. ALC 2015 specifically requires leaders to "be adept at framing complex, ill-defined problems through design and make effective decisions with less than perfect information" (U.S. Army, 2011). The ALC 2015 acknowledges the need to focus on the fundamentals that contribute to mission success.

> Mastering and sustaining core fundamental competencies better support operational adaptability than attempting to prepare for every possibility. The fundamental competencies must be clearly identified to support executing future full-spectrum operations and time must be allotted to *attain proficiency through repetition and time on task*. (U.S. Army, 2011)

The ALC 2015 describes the desired training capability to shift to individually-tailored instruction and take advantage of emerging learning technology capabilities. These capabilities include "Adaptive learning, intelligent tutoring, virtual and augmented reality simulations, increased automation and artificial intelligence simulation, and massively multiplayer online games (MMOG), among others will provide Soldiers with opportunities for engaging, relevant learning at any time and place" (U.S. Army, 2011).

> Adaptive learning and intelligent tutors. Technology-delivered instruction can adapt to the learner's experience to provide a tailored learning experience that leads to standardized outcomes. One-on-one tutoring is the most effective instructional method because it is highly tailored to the individual. While establishing universal one-on-one tutoring is impractical, the Defense Advanced Research Projects Agency (DARPA) and other research agencies are demonstrating significant learning gains using intelligent tutors that provide a similarly tailored learning experience. Through adaptive learning software, technology-delivered instruction adapts to the learner's previous knowledge level and progresses at a rate that presents an optimal degree of challenge while maintaining interest and motivation. Technology-delivered instruction that employs adaptive learning and intelligent tutoring could save time and allow for additional gains in learning effectiveness. (U.S. Army, 2011)

> Digitized learning content. Digitized learning content incorporates easily reconfigurable modules of video, game-based scenarios, digital tutors, and assessments tailored to learners. They incorporate the use of social media, MMOG, and emerging technologies. Interchangeable modules are easily shared and updated to stay relevant (U.S. Army, 2011)

In conclusion, the Army's FM 7-0, Training Units and Developing Leaders for Full Spectrum Operations, as well as the ground-breaking ALC 2015 creates a tremendous opportunity to develop and integrate game-based training tools to support critical training with improved results. However, ensuring that the training tools and scenarios developed meet the desired training objectives needs to be explored.

D. LITERATURE REVIEW

1. Learning and Educational Models

Many leader tasks and competencies within the Army are not well suited to the typical didactic learning that is so prevalent within the Army education institutions. The often-used Confucius quote, “Tell me, and I will forget, Show me, and I may remember, Involve me, and I will understand” directly applies to the game-based learning and the experiential learning model.

a. Constructivist Learning Environment

Wilson describes a constructivist learning environment as a learning environment that emphasizes “meaningful, authentic activities that help the learner to construct understandings and develop skills relevant to problem solving“ (1996). The foundation of the constructivist learning theory is that the student learns through concrete experiences that allow the student to put ideas to practice in a way that enables deeper understanding of relationships in nature (Jonassen, 1999). These relationships may not be well understood through didactic instruction as the only means of instruction due to the complexity of the relationships.

Wilson (1996) describes a learning environment as a “place where learners may work together and support each other as they use a variety of tools and information resources in their guided pursuit of learning goals and problem-solving activities.” Wilson then continues to describe the learning environment to include many environments to include computer micro-worlds.

The constructivist learning environment has seven pedalogical goals (Wilson, 1996):

1. Provide experience with the knowledge construction process where students take responsibility for strategies and methods for solving problems.

2. Provide experience in and appreciation for multiple perspectives where students are exposed to multiple acceptable solutions to enhance their own understanding of the problem.
3. Provide experience in realistic and relevant contexts where students are not able to isolate the tasks from outside noise.
4. Encourage ownership in the process where students are not able to take a passive role in their education and are required to make decisions.
5. Embed learning in a social experience where students influence and are influenced by other students.
6. Encourage the use of multiple modes of representation where students are responsible for representing their knowledge through several means.
7. Encourage self-awareness of the knowledge construction process where students are encouraged to not only know something, but are able to articulate how and why they know something.

Critics of the constructivist learning environment point to the challenge that it is difficult to ensure that all students will achieve the same learning outcome (Savery & Duffy, 1998). To prevent this undesirable outcome would require careful analysis of the learning environment to ensure the wrong things are not accidentally learned during the experience. The learning environment, like any game, model, or simulation, is an approximation of reality. It is important to ensure that critical components of the environment are appropriately represented and trivial components of the environment are minimized.

b. Experiential Learning Models

The experiential learning model is a method of education that seeks to provide students with a semi-structured educational environment where the subjectivity of the learning experience is understood. The experiential learning model uses exercises and experiences as the primary means of student learning. There are two primary models for the experiential learning model. The three stages of this model are “plan, do, and review.” This approach was developed by Dewey, who emphasized that student learning is the greatest when the students are actively engaged with student-directed education (Neill, 2012). In 1938, there was an educational debate (that continues today) between two schools of thought, which are: 1) relatively structured, disciplined, ordered, didactic tradition education, and 2) relatively unstructured, free, student-directed progressive education. Critics of the traditional educational model say that rote memorization of rules and ideas does not mean that the student understands how to apply them to the real world. The objective of education is not simply to memorize rules, but rather be able to apply knowledge to situations for an improved result. Critics of the experiential learning model are concerned that student-directed learning will not ensure that the students will ultimately learn the desired material.

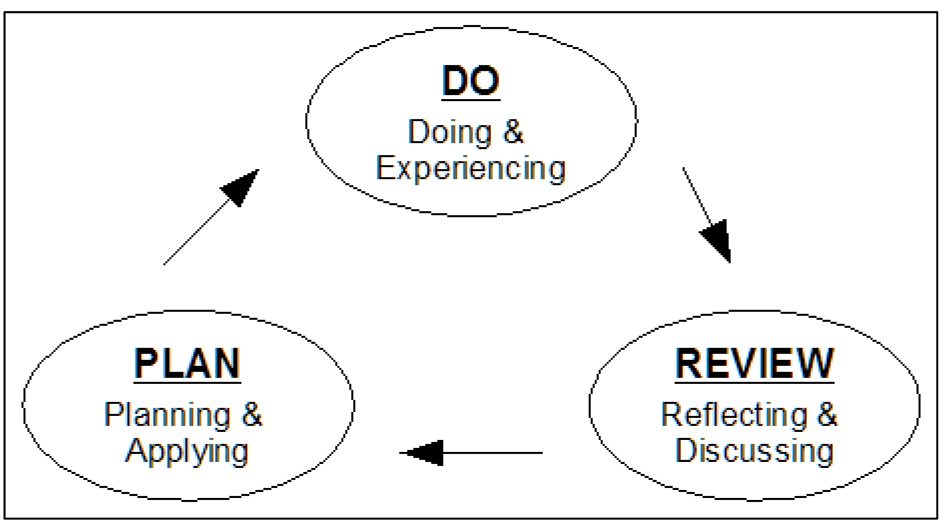

Figure 3. Three Stage Model of Experiential Learning (From Neill, 2012)

Kolb, in 1984, developed the “Experiential Learning Model” based on the previous model by Dewey. Kolb’s model is used in training and education communities today. The four stages are: 1) a concrete experience, 2) reflective observation, 3) abstract conceptualization, and 4) active experimentation. Exeter, in 2001, essentially re-used Kolb’s model, but added a “transfer of learning” component to the model (Neill, 2012). This transfer of learning addressed the previous concern about what students were ultimately learning from the experience.

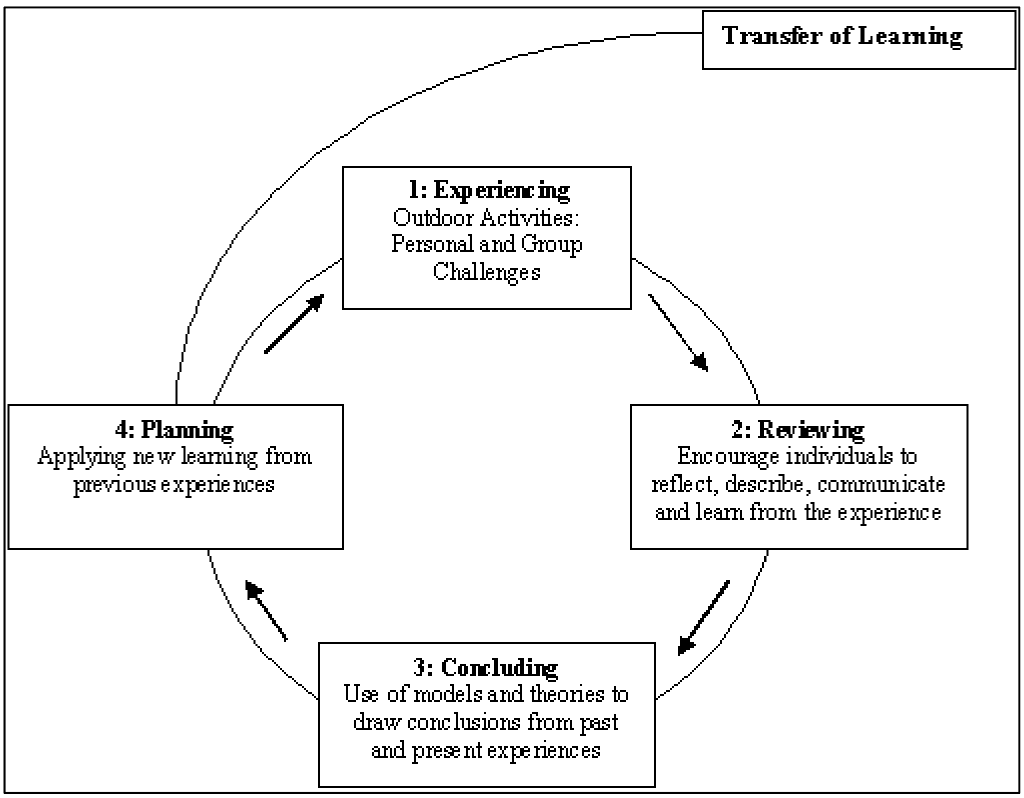

Figure 4. Four Stage Experiential Learning Cycle (From Neill, 2012)

In summary, the experiential learning model/cycle seeks to provide a higher quality of education to the student than just didactic methods. Game-based learning brings a unique attribute to address the concerns that you cannot be certain what the student learns in the experiential learning model. Game-based education can provide the student with a directed practice and experimental learning environment and yet control the learning by rewarding good performance and penalizing poor performance. These rewards and penalties reflect the desired learning objectives when done correctly. Game-based training provides the learning environment, but a evaluation method of the game and scenario is needed to provide verification for the training developer.

c. Ericsson's Deliberate Practice

In 1993, Ericsson, Krampe and Tesch-Romer described the role that deliberate practice had in the development of expert performance (1993). First, Ericsson et al., asserted that "sufficient amount of experience or practice leads to maximal performance appears incorrect" (1993). They found characteristics most effective in improving performance. First, students should receive immediate feedback and knowledge of results of their performance and the students should repeatedly perform the same or similar tasks. Second, to ensure effective learning, subjects should be given explicit instruction about the best method to perform the desired task and should be supervised by an instructor to allow individualized diagnosis of errors, feedback, and remedial part training. Deliberate practice is teacher designed practice activities that the individual engages in between meetings with the teacher (Ericsson, Krampe, & Tesch-Romer, 1993).

Deliberate practice is different from work and play. Ericsson et al., characterize ''work'' as directly motivated by external rewards and ''play'' is characterized as having no explicit goal and is inherently enjoyable (1993). Ericsson et al., state that deliberate practice includes activities that have been specially designed to improve the current level of performance (Ericsson, Krampe, & Tesch-Romer, 1993). Therefore, deliberate practice seeks to combine some of the characteristics of ''work'' and ''play'' to create an environment where the student is able to practice specified tasks repetitively in a low-cost and low-risk environment that provides an intrinsic reward that also provides focused feedback on learning objectives. Table 1 articulates the distinct differences between work, deliberate practice, and play as discussed by Ericsson et al.

Table 1. Difference and similarities between work, deliberate practice, and play adopted from Ericsson (After Ericsson, Krampe, & Tesch-Romer, 1993)

	Work	**Deliberate Practice**	**Play**
Tasks / Structure	Comprehensive – structured to meet real requirement	Part task or full task – structured specifically for the student	No structure
Reward	Extrinsic	Intrinsic	Intrinsic / Enjoyment
Repetitions	Limited	High	High
Feedback	Limited – typically outcome focused	Focused on learning objectives – process and/or outcome focused	Not typically used
Cost of mistakes	High	Low	None

There is an identified challenge with the current Army model for educating and training officers. Army leaders undergo supervised activities while learning the basic concepts in an institutional environment before arriving at a unit where they are expected to have a level of proficiency of the basic concepts. Then when the leader arrives to the operational unit, they are expected to give their best performance each and every time performing the tasks, which relies on previously learned methods rather than exploring alternative methods with undetermined consequences. Leaders understand that making mistakes is a critical part of training and education, but there are not enough resources such as time, money, and materials, to repeat the exercises enough to become proficient.

Therefore, there is great expectation for the leaders to perform at their best each and every time they conduct an exercise, which contradicts one of principles of deliberate practice.

Deliberate practice supports the vision of FM 7-0 and supports the guidance of the Army Learning Model 2015. To enable deliberate practice within the institutional, operational, and self-development domains, the Army is adopting games as a time and cost effective addition to the existing Live, Virtual, and Constructive simulations. These games provide an environment for leaders to practice their craft without the same level of resource expenditure of time, money, and materiel.

d. Performance Feedback

James Ong stated that "Practice and experience, whether simulated or on the job, are not enough to ensure effective learning. Learners must be able to make sense of those experiences to identify poor decisions and actions, missing knowledge, and weak skills that deserve attention" (2007).

Perhaps the most critical component of deliberate practice is performance feedback. Performance feedback encompasses more than just a message that you completed the exercise successfully. Performance feedback includes everything the learner perceives that helps them make connections between their actions (cause) and the outcome of those actions (effect).

There are many ways to provide performance feedback to the student during and after an exercise to influence learning. For well-defined problems, the tree diagram in Figure 5 describes the notion that games, as well as all training and education, should reward good performance and penalize poor performance and there are negative consequences to rewarding poor performance and penalizing good performance.

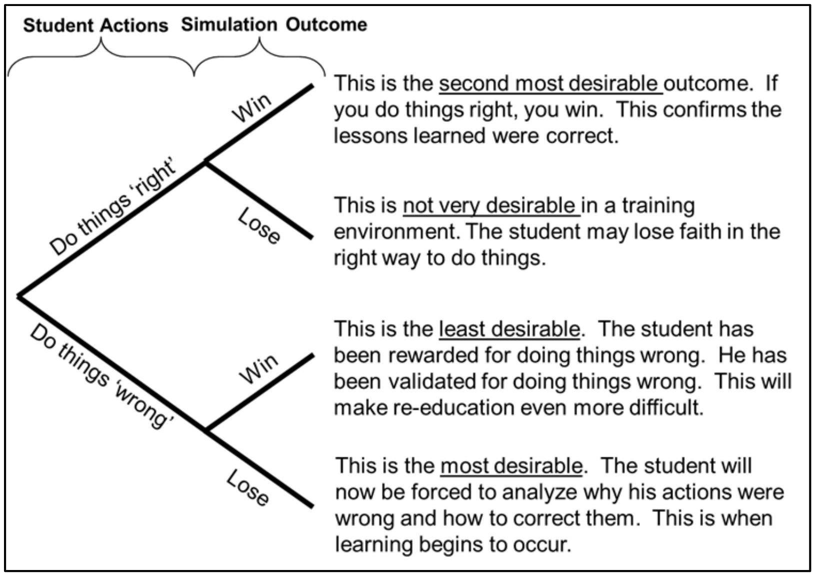

Figure 5. Performance Feedback Tree Diagram for Well-Defined Problems.

This tree diagram can also be represented in a matrix that is analogous to statistical Type I and Type II errors, where Type I error is analogous to providing negative feedback for correct performance, and Type II error is analogous to providing positive feedback for incorrect performance.

		Performance Feedback	
		Reward	Penalty
Student Performance	Correct	Desirable	Not Desirable – student received negative reinforcement feedback from correct performance
	Incorrect	Not Desirable – student received positive reinforcement feedback from incorrect performance	Desirable

Figure 6. Performance Feedback matrix for well-defined problems.

Performance feedback for ill-defined problems is not as straight forward as it is for well-defined problems. Clark describes ill-defined tasks and problems as “scenarios or cases for which there is no one correct answer or approach… ill-structured problems are considered best for problem based learning” (Clark, 2008). Ill-defined problems are also characterized as problems where there exists a range of acceptable solutions and a range of unacceptable solutions. In the range of acceptable solutions, the solutions may be very different from each other, but still adequately address the problem and should be rewarded equally. Figure 7 graphically depicts this notion as it relates to performance feedback.

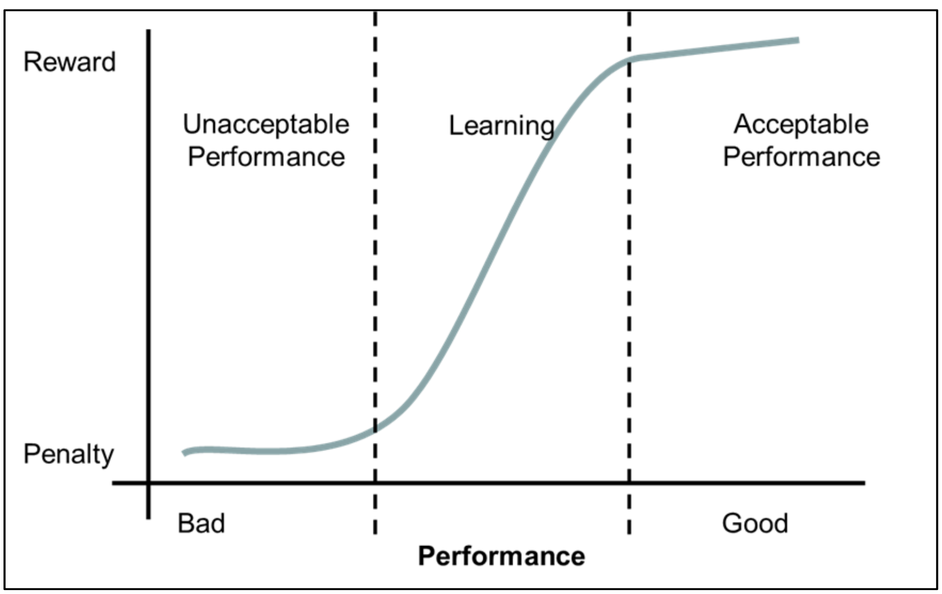

Figure 7. Reward function as it relates to performance.

The ‘‘unacceptable performance’’ region of this curve refers to performance that is unacceptable and is used to identify students that do not have a requisite knowledge to begin deliberate practice. The learning portion of the curve is very important for student learning. This region is where students depend on the reward associated with their performance to gain insights about which strategy is better than other strategies. The acceptable performance region indicates where student performance matches the desired training or educational goals of the exercise. This curve is utilized, in practice, in the entertainment game industry to keep players in what Murphy refers to as ‘‘flow’’ or the learning portion of the curve. (Murphy, 2011) This supports the intrinsic rewards found in play by Ericsson.

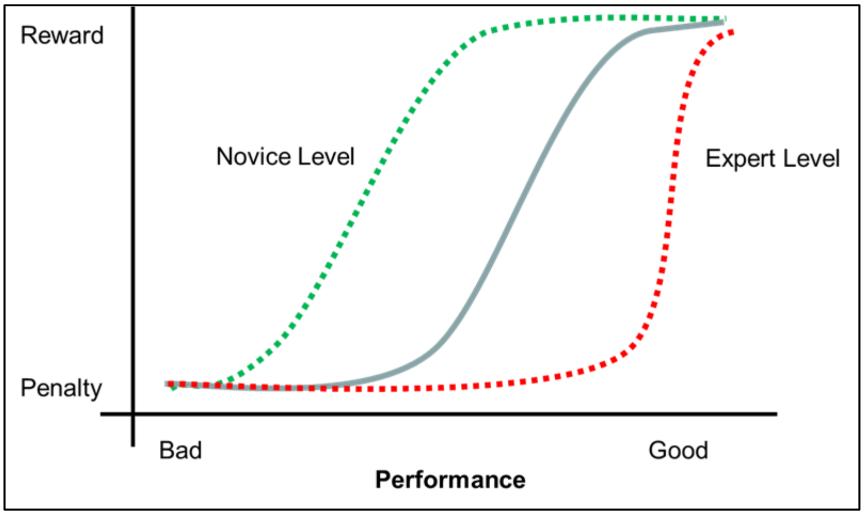

Figure 8. Manipulation of reward curve for games.

The reward function curves can also be used to evaluate existing training simulations and scenarios. The following charts show a few hypothetical reward functions that do not support the desired training objectives. Figure 9 describes a reward function that rewards mediocre performance over good performance. This is undesirable because students would perceive their mediocre performance as the desired good performance.

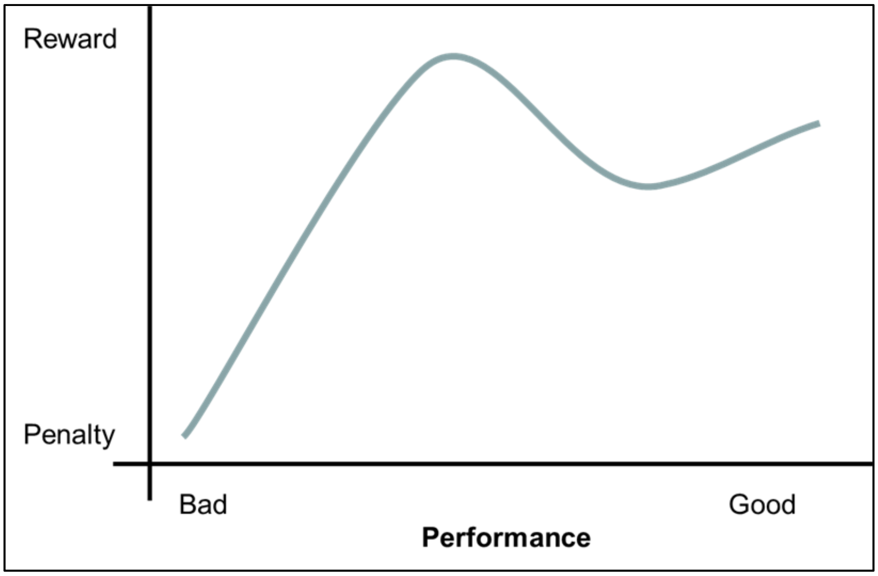

Figure 9. Undesirable reward function that rewards mediocre performance

Figure 10 describes a reward function that does not adequately differentiate good performance from bad performance. This is undesirable because students perceive that there is no way to ''win'' and no way to ''lose'' so they do not adjust or improve their performance to obtain good performance.

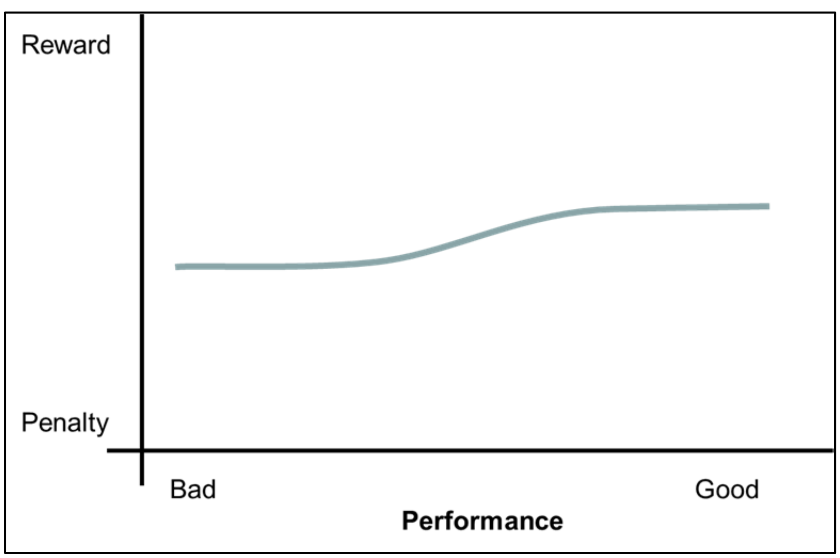

Figure 10. Undesirable reward function curve that does not adequately differentiate between good and bad performance.

2. Game-Based Learning

Games are different from simulations in a few significant ways. First, simulations seek to model a potential event, phenomenon, or outcome that occurred, or could occur in the real world. Games are different in that they focus on the experience of the user or player. Game developers seek to use plausible simulation data to drive the outcomes of events, but developers will modify the outcomes of the simulation to meet the entertainment needs of the game (Kapp, 2012). Traditionally, games have been used exclusively for entertainment. However, there have been many cases where things learned in the game environment have had applicability in the real environment (Fullerton, 2008). Therefore, the outcomes of the events in the game do not necessarily need to represent reality, but they must entertain the player. When games are used for training, once again, the outcomes do not have to represent reality, but they must educate or train the user appropriately for the game to be successful.

The second way that games are significantly different from simulations is the use of a reward signal. Simulations seek to model a potential event, phenomenon, or outcome that occurred, or could occur in the real world. Simulations do not explicitly provide a reward signal for the user. Simulations can provide the stimulus for the user to determine a reward. For example, in a simulation, a student positions a force in a concealed fighting position and the unit successfully defends the position from an attack. The next time the student places the force in the open without any concealment and the unit does not successfully defend the position from attack. The student could construe that he perceived a reward by using concealment and this would be accurate. However, students would have to provide their own goal or objective in order to perceive this reward. A game explicitly states the goal or objective for the student.

3. Current Games Used for Tactical Military Training and Education

a. Command Mentoring Intelligent Tutoring System

Command Mentoring Intelligent Tutoring System (ComMentor), developed by Stottler-Henke Associates, is an experimental effort sponsored by the Army Research Institute, which emulates the Socratic teaching methods used by expert instructors. ComMentor presents tactical scenarios of major combat operations to students and prompts them to enter their responses via graphical user interfaces, form-structured text, and tactical maps. As with ill-defined problems, there is no single correct answer to a scenario, so ComMentor evaluates each student's reasoning skills by comparing their solutions and rationale with fragments characterizing expected appropriate and inappropriate student responses supplied by experts. ComMentor uses these assessments, along with structured arguments, to control its line of Socratic questioning, hinting, and feedback to enhance the student's high-level thinking habits (Stottler-Henke Associates, 2012).

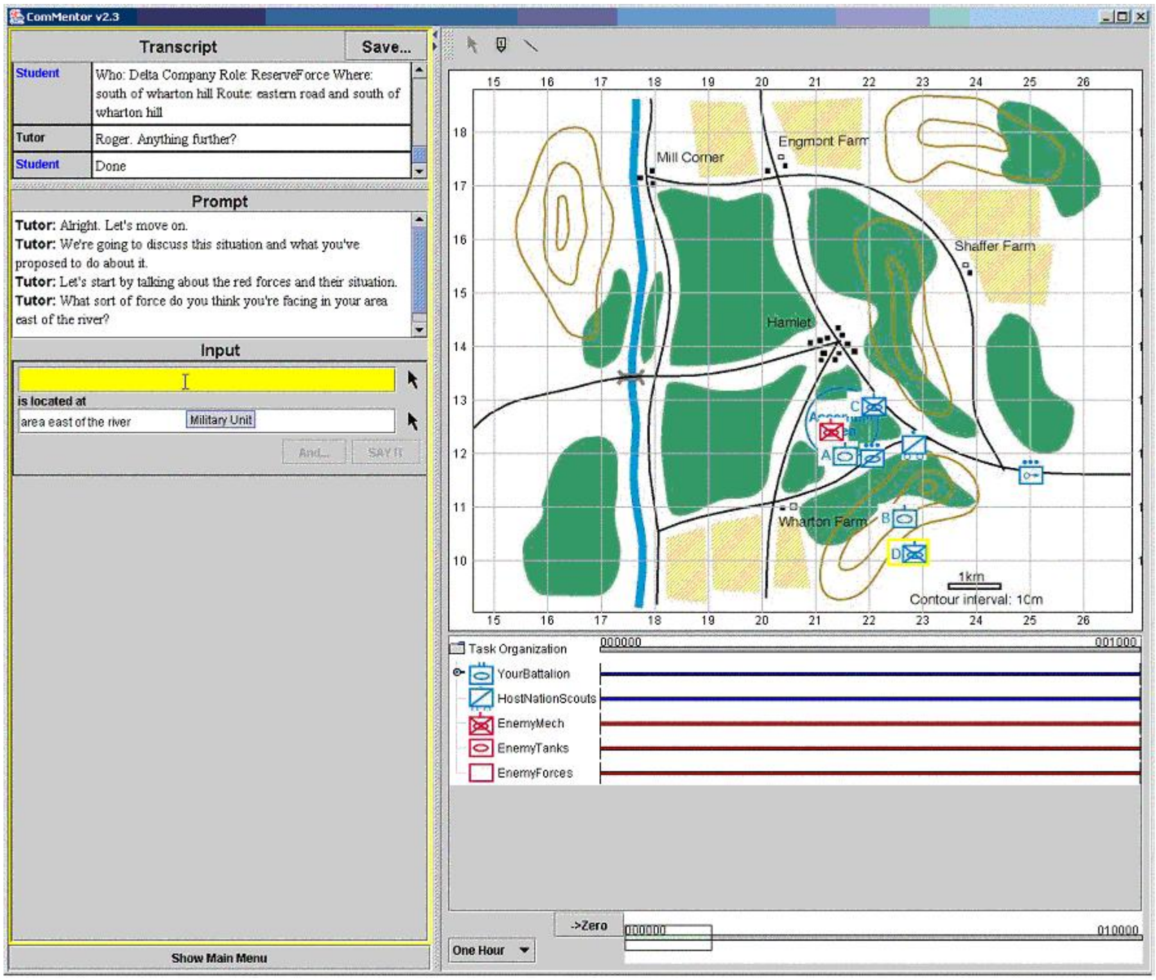

Figure 11. Command Mentoring Intelligent Tutoring System (ComMentor) interface

ComMentor, from an intelligent tutoring system perspective, sought to instruct students on the process of decision-making as well the execution of the decisions. The outcome of decisions were scripted to meet the education objectives and is not (Stottler, Jensen, Pike, & Bingham, 2002) an open-ended simulation. The primary means of interaction in ComMentor is the Socratic dialogue that is scripted by subject matter experts prior to the exercise. The effort to develop a training scenario with the included authoring tools is approximated to be “14–20 days—roughly 1 person-month of effort” (Domeshek, Holman, & Luperfoy, 2004) In addition to time, it is estimated that authoring a scenario would cost $50,000 per scenario developed by skilled personnel. (Domeshek, Holman, & Luperfoy, 2004) Due to the high reliance on the scripted interaction with the student, the student is presented with a tactical situation, makes decisions, discusses decisions with scripted tutor, is coached to the proper solution, and then is presented with the next tactical situation. While this Socratic interaction has a positive impact on the student's learning, it does not allow the student to deal with the negative (or positive) consequences of their decisions. It is similar in nature to a golf scramble. Everyone tees off and the best ball is played by all of the players. If you hit it into the woods, you do not have to play it out of the woods. In ComMentor, if you make a tactical error, you do not have to fight through the consequences of that decision, but rather you are coached to the right solution before you go on to the next situation. For the Socratic interaction to work properly, the expert developing the scenario must appropriately anticipate the entire range of potential student solutions to the particular tactical situation. This necessitates limiting the potential student solutions to the tactical situation. Through the Socratic interaction, the student will change his course of action to align with the instructor-desired course of action before the next tactical situation is presented. This structure for the exercise does not lend itself to students repeating the exercise or exploring other potential solutions because of significantly diminished returns executing the same exercise with the same feedback more than once. Therefore, the scenario, which is rather expensive, is designed for the student to execute once and limits the reuse capability.

The Army Research Institute (ARI) sponsored research found that the Socratic intelligent tutoring system was effective, however, required significant resources to develop. It cost roughly $50,000 to develop each scenario and required over 100 hours of dedicated subject matter expert involvement (Domeshek E. , Technical Report 1124 Phase II Final Report on an Intelligent Tutoring System for Teaching Battleifield Command Reasoning Skills, 2004). As a prototype, users found that ComMentor had a limited range of choices or options available for the learner. This shortcoming can prevent the learner from exploring many potential solutions and is desired in the experiential learning model.

b. Battle Command 2010

Battle Command 2010 (BC2010) is a tactical decision game designed by Mak Technologies with an Intelligent Tutoring System developed by Stottler Henke Associates. (Stottler Henke Associates, 2012)

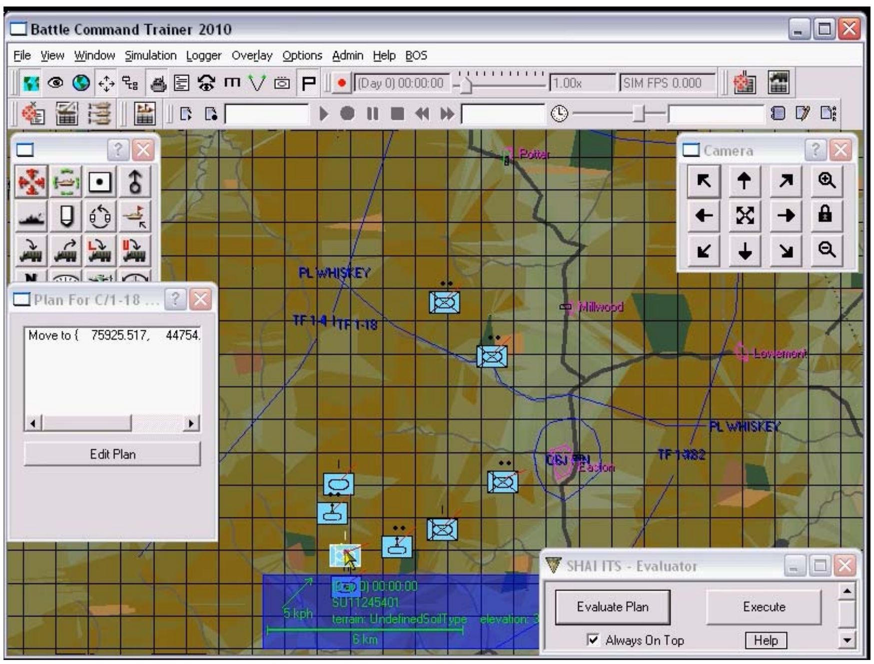

Figure 12. Battle Command 2010 (BC2010) Interface

BC2010 is based on a tactical simulation so that the students are able to experience the consequences of their decisions. The tactical simulation adjudicates the interaction between opposing forces and displays the results for the player to make a decision. These interactions are not pre-defined by the scenario author but are the result of free-play. Therefore, the performance feedback mechanisms depend on observable accomplishment of certain simulation states that involve unit location and actions. The intelligent tutoring aspect of this training system requires the student to select ‘‘evaluate this plan’’ button on the graphical user interface. Selecting ‘‘evaluate this plan’’ causes an algorithm to run that compares the student’s performance to pre-generated instructor feedback and displays the appropriate feedback. For example, the instructor suspects that students may wrongly choose course of action A, so the instructor prepares specific feedback to address the mistakes made when selecting course of action A. If the student during the exercise chooses actions similar to course of action A, the game will display the specific feedback the instructor prepared while authoring the scenario.

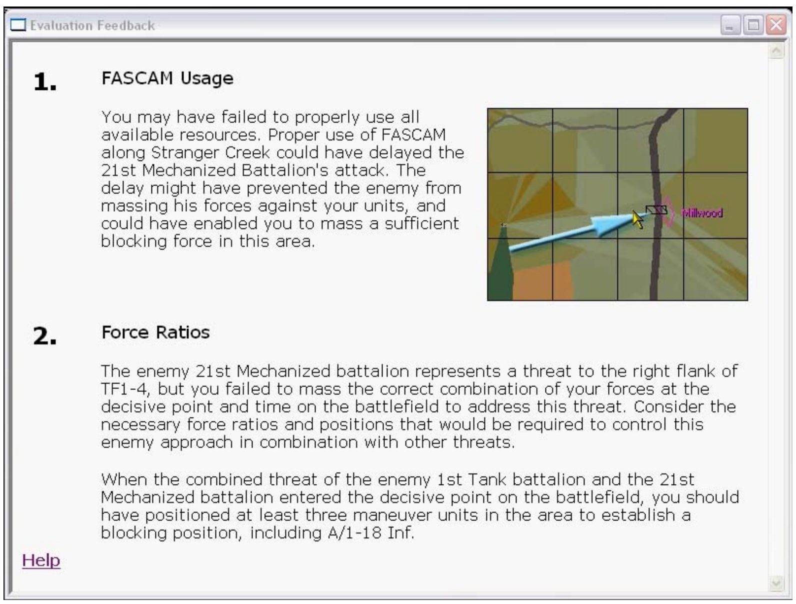

Figure 13. Evaluation Feedback from BC2010.

The performance feedback is based on the student’s decisions, but similar to ComMentor, the expert must anticipate the student’s actions when authoring the scenario. Additionally, this supposes that there is a single correct solution to the tactical situation. The feedback is not tied to the outcome of the decisions, but the decision itself. This can be problematic when the student pre-empts an enemy action that negates reactive actions later, however, the tutoring system is still looking for the reactive decision that is inconsequential. The free-play aspect of this training system facilitates repetition, however, it is limited due to the fact that enemy actions are scripted and do not change with each iteration (Stottler, Jensen, Pike, & Bingham, 2002).

c. Tactical Action Officer Intelligent Tutoring System (TAO ITS)

Stottler-Henke Associates developed the Tactical Action Officer Intelligent Tutoring System (TAO ITS) to support the Surface Warfare Officer School. Stottler stated that, "Experts and instructors agree that the most important factor for maintaining a TAO''s tactical decision-making skill is the opportunity to practice making decisions and timely feedback" (Stottler & Vinkavich, Tactical Action Officer Intelligent Tutoring System (TAO ITS), 2000). This observation is consistent with Ericsson's deliberate practice model. The TAO ITS displays realistic scenarios for the Tactical Action Officer (TAO) to observe, understand, and make a decision about what to do in the particular situation. If the students do not do the right things in the scenario, the students are faced with the consequences of their decisions.

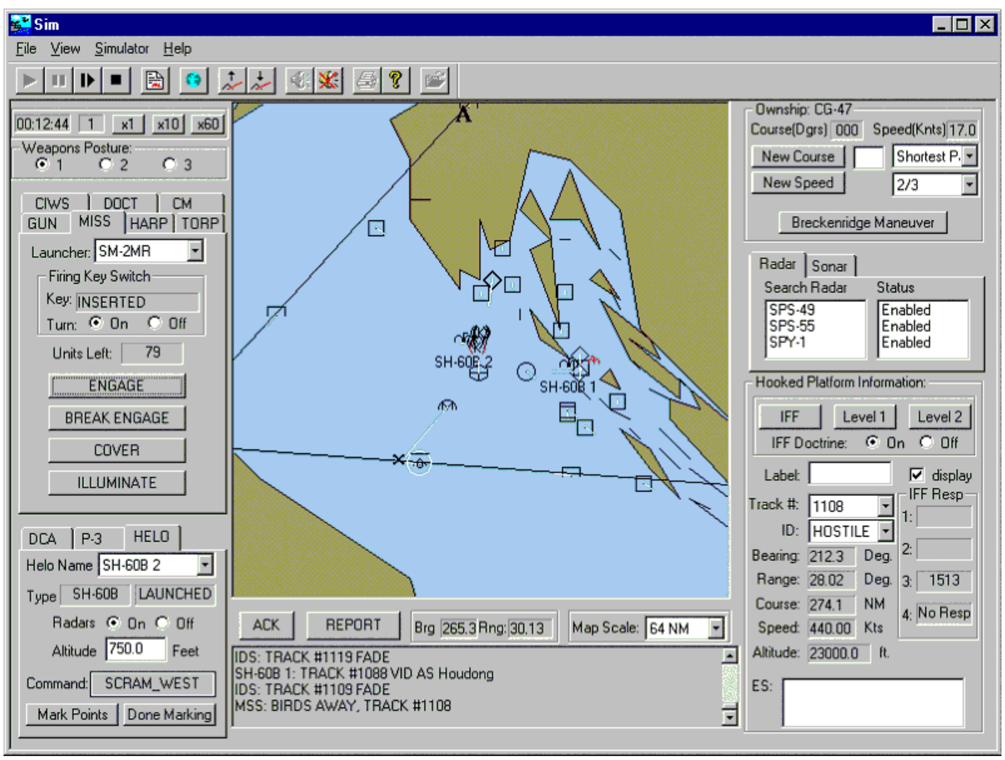

Figure 14. The Navy''s Tactical Action Officer Intelligent Tutoring System (TAO ITS)

The TAO ITS creates a student file for each student and tracks their performance of tasks through multiple exercises. This facilitates the instructor to give exercises that focus on identified student shortcomings. This capability supports Ericsson's deliberate practice model where each deliberate practice is structured to meet the needs of the student.

Following the TAO ITS exercise, the student is presented with performance feedback. This feedback is indexed to the exact time the student made, or did not, make a decision. This enables the student to see what input they observed, their decision, and the ''correct'' decision at that particular time in the exercise. This knowledge of performance and feedback enables improved performance. The student is able to repeat the exercise to perform the tasks correctly, however, there are diminished returns from repeating the exercise more than a few times because the scenario is scripted. Therefore, after a few iterations of the exercise the student is not reacting to the stimulus of the exercise, but rather making decisions based on what they know to be the correct answer at the particular time in the exercise.

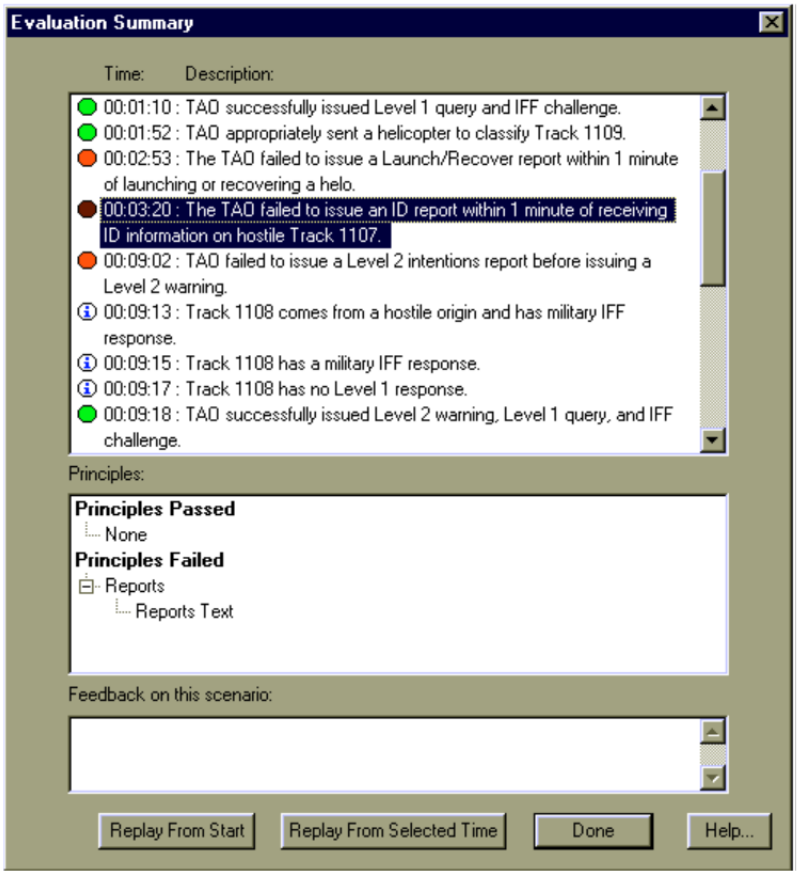

Figure 15. TAO ITS Performance Feedback (From Stottler & Vinkavich, Tactical Action Officer Intelligent Tutoring System (TAO ITS), 2000)

The Surface Warfare Officer School use of TAO ITS has improved the ability of Navy surface warfare officers to achieve significantly higher scores on standardized tests and student confidence has improved (Stottler & Vinkavich, Tactical Action Officer Intelligent Tutoring System (TAO ITS), 2000).

E. URBANSIM AND PSYCHSIM

1. UrbanSim

The U.S. Army directed Research, Development, Engineering Command (RDECOM) Simulation Training Technology Center (STTC) to develop a desktop tool that would support education and training objectives associated with counterinsurgency operations that the Army was having difficulty with in Iraq and

Afghanistan. RDECOM STTC worked closely with University of Southern California (USC) Institute for Creative Technology (ICT) to develop a game to address the unique challenges battalion and brigade commanders were facing in Iraq (McAlinden, Durlach, Lane, Gordon, & Hart, 2008). The development team interviewed returning battalion and brigade commanders to understand the type of challenges they were faced with during their time in Iraq. Following these individual interviews, the team collated the information and presented it to the recently formed counterinsurgency academy at Fort Riley, Kansas as well as the Combined Arms Center at Fort Leavenworth, Kansas to ensure their understanding was consistent with current doctrine and recent lessons learned from Iraq. Next, the development team developed UrbanSim reusing a previously developed piece of software called PsychSim to adjudicate the changes to the game environment and, by extension, provide feedback to the learner. After the game was developed, it was tested to ensure stability on the intended computers and fielded to the School for Command Preparation (Wansbury, 2011). Play testing was limited to ensuring functionality. The development team then waited for comments and concerns from the users about any problems they encountered with the system or within the game-play. Only a few problems were identified and those problems have been addressed by subsequent versions of UrbanSim.

UrbanSim was originally intended to be used at the School for Command Preparation to prepare Lieutenant Colonels and Colonels to command battalions and brigades. However, the UrbanSim package spread to other schools and institutions within the Army. Currently, UrbanSim is being used for instruction at:

- School for Command Preparation (SCP), Fort Leavenworth, Kansas—Army Lieutenant Colonels and Colonels preparing to command battalions and brigades
-

- Intermediate Level Education (ILE), Fort Leavenworth, Kansas—Army Majors preparing to serve as battalion operations officers, battalion executive officers, and other battalion and brigade staff positions
- Maneuver Captain's Career Course (MC3), Fort Benning, GA—Army Captains preparing to command infantry and armor companies and serve on battalion and brigade staffs
- Maneuver Support Captain's Career Course (MSCCC), Fort Leonard Wood, MO—Army Captains preparing to command combat engineer companies and serve on battalion and brigade staffs
- Warrior Skills Training Center, Fort Hood, TX—Army Non-commissioned officers (NCOs) preparing to serve in a large variety of leadership positions from the squad to battalion level

Currently, UrbanSim and several scenarios are available to the entire Army through the Military Gaming website. This enables all soldiers and leaders to access this software training tool for individual professional development.

UrbanSim supports experiential learning in ways that previous efforts with ITS can not achieve. Many of the other ITS are constrained by the scenario author anticipating student decisions during the design process. UrbanSim provides a rich environment for users to perceive the cause and effect relationship of their decisions in the environment. However, to achieve the desired training capability described in ALC 2015, and supported by learning science, a means to evaluate the performance feedback mechanism is needed for UrbanSim.

2. PsychSim

PsychSim is a social simulation tool for modeling a diverse set of entities (e.g., people, groups, structures), each with its own goals, private beliefs, and mental models about other entities. Each agent generates its beliefs and behavior by solving a observable Markov decision problem (Wang et al., 2012) PsychSim has been used in other fielded Army simulations and games for training and education. Elect BiLAT utilizes PsychSim as the underlying simulation to adjudicate the interaction between the player and an avatar that represents a key leader in a controlled cultural context.

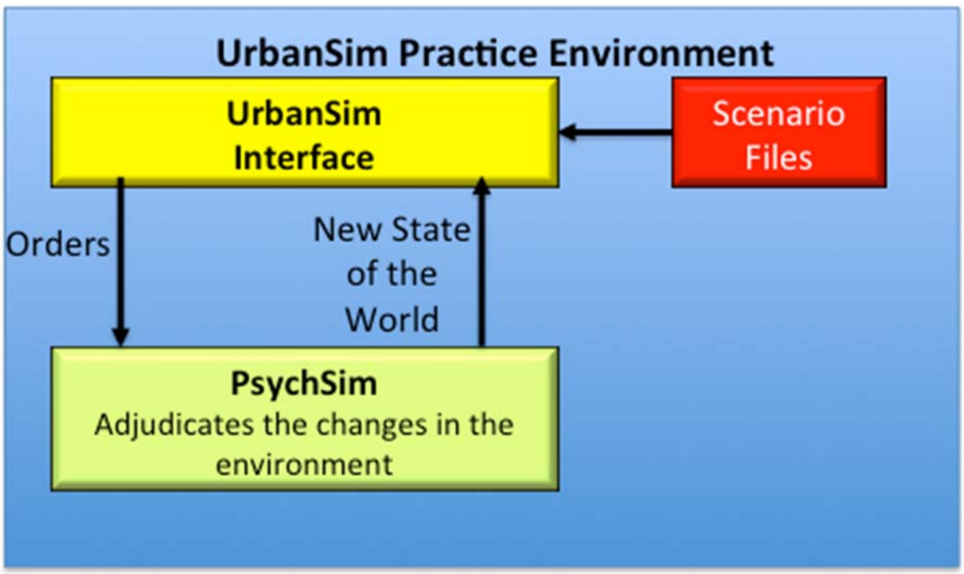

Figure 16. UrbanSim Practice Environment - UrbanSim/PsychSim relationship

3. UrbanSim Performance Feedback Mechanisms

There are several ways that the player receives feedback during the game play. This study focused on the Lines of Effort assessment at the primary means of performance feedback to the student.

a. Lines of Effort (LOE) Assessment

During game play, the student is able to view the current status of six lines of effort. The lines of effort are on a 0 to 100 scale, and are Civil Security, Governance, Host Nation Security Forces, Essential Services, Information Operations, and Economics. Following each turn the LOE is updated along with a red or green arrow to denote an increase or decrease in that particular LOE.

Figure 17. UrbanSim Interface Line of Effort feedback

b. Population Support Meter

The other performance feedback indicator that is always present on the graphical user interface is the population support meter. The population support meter represents the percentage of the population that supports our efforts, is neutral to our efforts, and against our efforts.

Figure 18. UrbanSim Interface - Population Support Meter

The population support meter has been found by users to be rather unreliable as a measure of performance (Wansbury, 2011). There are circumstances where the LOEs improve but the population support meter does not. This is an example of contradictory performance feedback, which also violates the principle of appropriate performance feedback as a part of deliberate practice.

c. S2 and S3 Recommendations

After each turn, there is occasional feedback and recommendations from a notional S2, Intelligence Officer, and a notional S3, Operations Officer. This feedback is scripted during scenario generation and displayed if certain conditions exist during the game.

d. Analysis Feedback

UrbanSim provides some analytic feedback that can be used to better understand the cause-effect relationship between actions in the game environment and the student’s decisions. The primary analytic tool is the trend analysis.

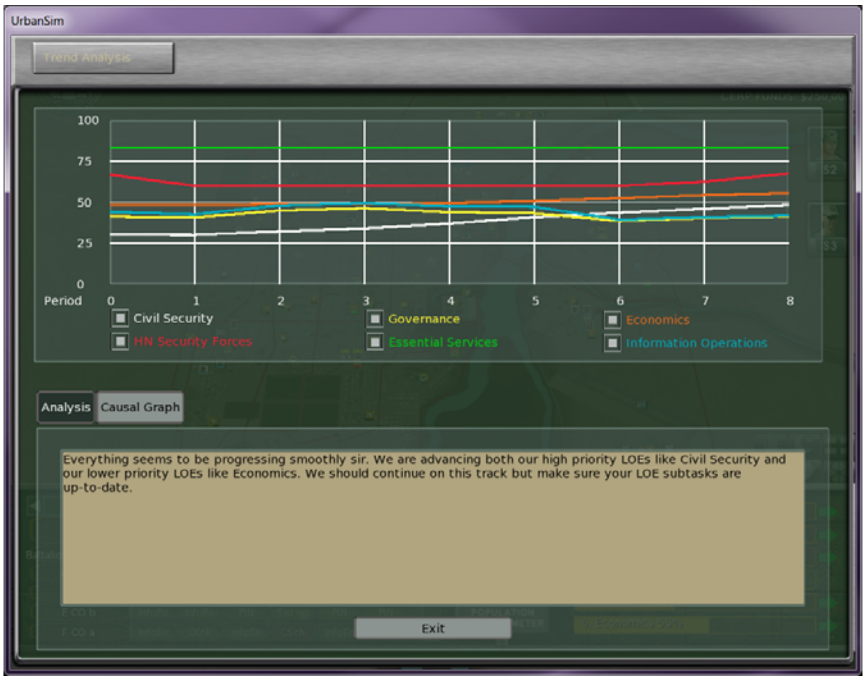

Figure 19. Trend Analysis within UrbanSim

The trend analysis shows how the various LOEs changed over the course of the game. This analysis is further refined for the user with the addition of a causal graph. The causal graph depicts the actions, results and how it changed the LOE. Red lines between the blocks indicates a negative result, and a green line indicates a positive result. It is possible for the same action to negatively affect one LOE, but positively impact a different LOE.

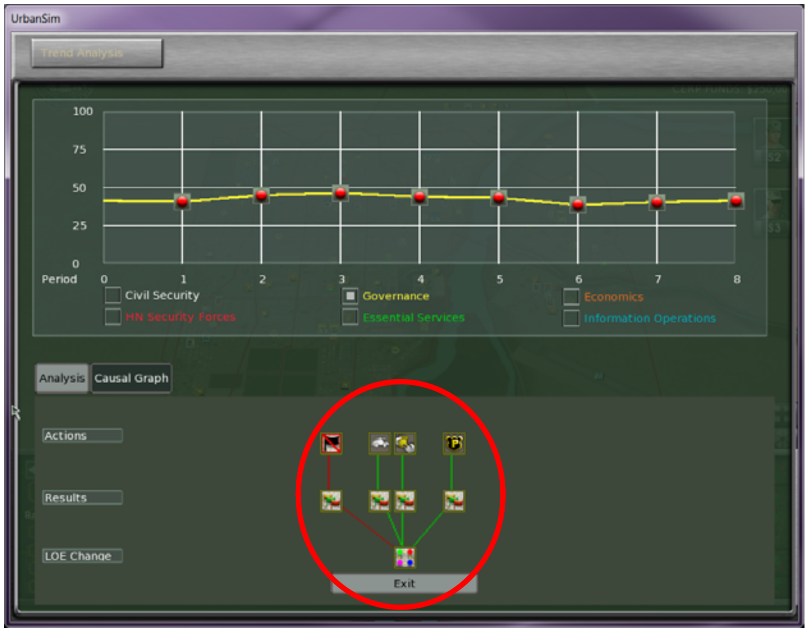

Figure 20. The Causal Graph with the trend analysis of UrbanSim

Within the trend analysis interface, there is a tab that takes the user to a causal graph that explicitly portrays why a particular LOE was affected in a particular turn. The presentation is well organized with the actions portrayed on top of the graph which are linked to results with red and green lines for positive and negative impacts respectively. The results are then connected to the LOE Change at the bottom. This enables the user to see how and why the LOEs changed in a particular turn. It is important to note that many of the actions described are not user decisions or actions, but rather actions the agents in the simulation autonomously do based on agent descriptions in the scenario file.

F. GAME PLAY TESTING

1. How Entertainment Games are Play Tested

Games that are designed for entertainment are play tested to ensure they meet both system requirements and well as providing entertainment to the player. Their focus is on the interaction between the real player and the game environment to ensure that it is entertaining and engaging. The primary use of automated play testing is to ensure software stability and to confirm that there is not anything the player can input that would cause the system to crash unexpectedly. Since, games are focused on the human entertainment value, the primary means of game play testing is with human focus groups representing the population they expect would play the game. These tests are resource intensive in terms of time and money.

2. How UrbanSim Developers Recommend Developing and Play Testing Scenarios for Training

> Play testing and balancing is critical to ensuring the scenario plays the way it is intended to and that it is as difficult or as easy as you the author or the training developer wants it to be. You should first play the scenario yourself a few times to make sure it is working the way you intended. It is highly recommended that you do this while building out the scenario instead of doing it at the end. This will allow you to spot problems early on and prevent headaches in the future.
>
> When your scenario is finished, play test to achieve every possible outcome in your scenario. This will give you a rough indication of whether the scenario is too difficult or too easy. You'll have to adjust the scenario accordingly to achieve the right level of difficulty.
>
> If possible, let other people play test the scenario and provide feedback. Because of your familiarity to the scenario, you will always have the advantage of "knowing too much" that other players will not when they play the scenario. The feedback that other players provide will be invaluable information as to whether your scenario is too difficult or too easy. Other players may also find problems in your scenario that you won't find by yourself. By play testing and balancing, you will provide the polish your work needs to better achieve the goals of your scenario. (U.S. Army RDECOM, 2011)

This description from the UrbanSim documentation about play testing is similar to the way that play testing is done for entertainment games. However, UrbanSim is intended to be a training game where the focus should be on ensuring that the desired player performance is rewarded and poor performance is penalized. Therefore, a different approach to play testing is needed to verify training games and scenarios.

3. Recent Efforts towards Automatic Verification of Training Simulations

Wang (and Pynadath and Marsella) recently published an article that describes an innovative way to playtest UrbanSim to determine whether the scenarios support the desired training objectives. Wang et al. point out that:

> From an instructional perspective, the use of complex multiagent virtual environments raises several concerns. The central question is what is the student learning—is it consistent with training doctrine and will it lead to improved student's performance? (Wang et al., 2012)

As training simulations and games for training become more prolific, increase in complexity, and provide deeper levels for student decisions, it becomes increasingly more problematic to verify the desired underlying pedagogy is present (Wang et al., 2012). Human play testing is a preferred method because of the accuracy of the results. However, as the complexity of the game increases, human play testing is only able to test a smaller portion of possible student strategies. Wang et al., concludes that, "Although multiagent systems support automatic exploration of many more paths than is possible with real people, the enormous space of possible simulation paths in any nontrivial training simulation prohibits an exhaustive exploration of all contingencies" (Wang et al., 2012).

Wang et al. conducted an experiment to determine the training impact of the training videos associated with the UrbanSim training package. The research team found that students that watched and implemented the ''Clear, Hold, Build'' strategy that is prescribed in both the videos and the Army's current doctrine performed better than students that did not view the videos. The research team developed and used Markov chain Monte Carlo (MCMC) simulation to develop a method for automated verification testing. They found that this method generated more incorrect strategies than when humans played the scenario, but the overall distribution of scores were similar to the scores from human players (Wang et al., 2012).

G. CONCEPTUAL MODELS OF CURRENT AND PROPOSED SCENARIO DEVELOPMENT MODELS

1. Current Education Game Scenario Development Model

Many games and scenario development methods follow the conceptual model in Figure 21. Starting from the training objectives, the scenario is developed. The scenario designer typically tests different components of scenario as an anecdotal formative test. Then the scenario is fielded to the intended users. If there are any identified problems with the scenario, they are collected and corrected as time and resources permit.

Figure 21. Current training and education game scenario development model

Occasionally, games and game scenarios are explicitly evaluated against the intended training objectives. This explicit evaluation is typically done through academic research efforts and not generally done in operational organizations. When explicit evaluation is conducted, it occurs after the scenario development is complete.

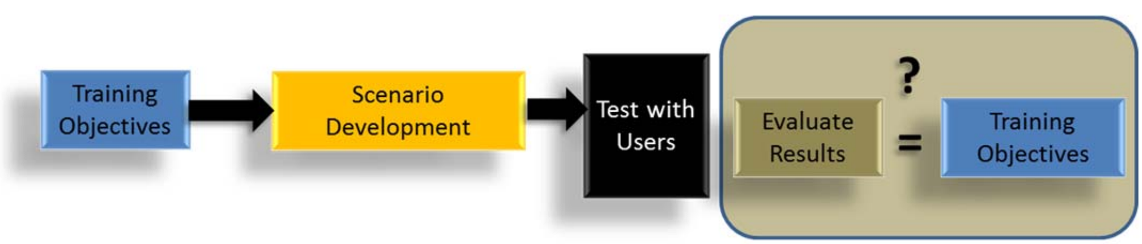

Figure 22. Game Scenario development model when training effectiveness is explicitly evaluated

2. Entertainment Game Scenario Development

Within the games for entertainment industry, there are many ways that games and scenarios are created and delivered to customers. However, they generally follow the pattern described in Figure 23.

Figure 23. Game scenario development model used in entertainment game industry

The scenario development starts with the game design objectives and includes human play testing. The results of the human play testing are compared to the game design objectives. If there is a mismatch, the design team goes back to the scenario development effort. When the results of the human play testing match the desired objectives of the game design, the game is delivered to customers.

3. Proposed Education Game Scenario Development Model

Using automated formative evaluation tools can facilitate a greater success rate of meeting the training objectives when play tested with humans or when directly fielded to the users. Figure 24 describes this proposed development model.

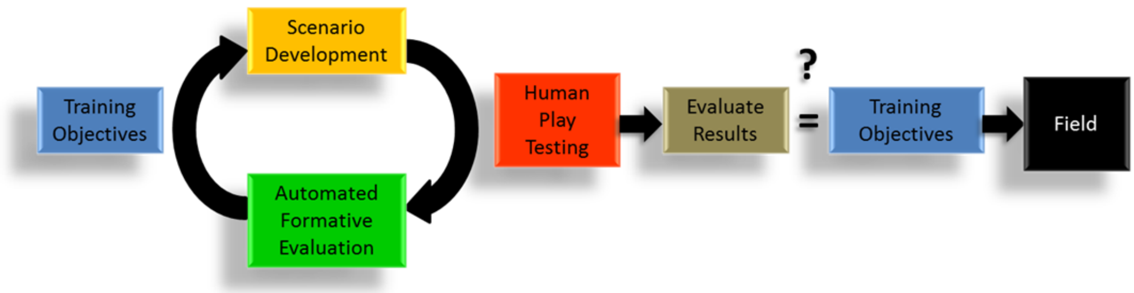

Figure 24. Proposed education game scenario development model using automated formative evaluation tools.

Similar to the previous models, scenario development starts with the training objectives. However, during scenario development, the designer uses automated formative evaluation tools to guide the development. This model follows the software development axiom of ‘‘build a little, test a little.’’ This allows for correction when the problems are relatively easy to identify and fix. Once this cycle is complete, human play testing is conducted to ensure the scenario meets the training objectives. The results of the human play testing are once again compared to the training objectives. The automated formative testing should provide more successful training objective achievement and reduce the amount of corrections needed after fielding.

The automated formative evaluation techniques are discussed and demonstrated in Chapters III and IV of this thesis.

H. REINFORCEMENT-LEARNING

Reinforcement-learning is a subfield of artificial intelligence based on behaviorist psychology. The goal in reinforcement-learning is to learn what action to take in a given situation in order to maximize long-term reward. The learning agent is tasked to learn the value of each action in a given state so that it can choose actions that provide greater value.

The components of a reinforcement-learning system are exploratory policy, reward function, and a value function (Sutton & Barto, 1998). These components are applied to an environment that has objects that interact with each other based on rules.

The exploratory policy describes how the agent will behave in a given time and situation (Sutton & Barto, 1998). For example, given a certain situation the exploratory policy describes how a particular choice is made by the agent. This can be similar to how a human player would act in a particular situation in a game.

The reward function describes the means the agent perceives the usefulness of particular actions (Sutton & Barto, 1998). The reinforcement-learning agent's sole objective is to maximize the reward in any particular situation and the reward function is used to assess how each action contributes to achieving the maximum reward. For games, the reward function may be the score, a particular outcome, or any quantifiable or qualitative observation of the environment. The reward function may include things that are out of the agent's control, but must be tied to the decisions made by the agent for learning to occur. For example, if the score of the game has no relation to the actions of the agent, or player, then no real learning can occur.

The value function is related to the reward function. While the reward function identifies what is good right now, the value function determines what is good in the long run (Sutton & Barto, 1998). The value function is used to determine the expected total reward the agent can accumulate in the future based on the current state. It is possible, and likely, that agents correctly choose an action that brings a lower reward in the short term because the value of that new state is higher than the value of choosing an action that brings a higher immediate reward but a much lower value. A simple analogy of this concept is people choosing to work at something unpleasant because they understand the long-term accumulation of rewards outweigh the current, temporary low reward.

Reinforcement-learning algorithms can be used to explore very large and complex decision spaces to provide insights about the underlying reward structure of a game or scenario. While identifying the greatest rewarded strategy is often the desired goal of using a reinforcement-learning algorithm, it also provides us a general ranking of the other possible strategies based on the perception of the learning agent.

The strength of using reinforcement-learning algorithms to explore large and complex decision spaces is that not all combinations of actions have to be tested or explored. Design of experiment techniques can also reduce the number runs of an experiment, but reinforcement-learning agents are able to dynamically assess and select policies during the experiment. Reinforcement-learning algorithms cannot guarantee an optimal solution in most applied cases, but can provide insight about the underlying reward structure. Reinforcement-learning algorithms are well suited for ill-structured problems and the evaluation of experiential learning platforms because the algorithm examines the scenario reward functions exclusively. This examination is the result of many more trials than are feasible with human players.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODOLOGY

A. METHODOLOGY TO EVALUATE GAMES AND SCENARIOS THAT ADDRESS ILL-STRUCTURED PROBLEMS

The following methodology was developed to evaluate UrbanSim scenarios for this research effort. However, this general methodology could be used, or adapted, to evaluate other games and scenarios.

1. Identify the training objectives. The training objectives are usually described in terms of what performance the learner should perceive a reward. However, it is equally important to understand what performance the learner should perceive a penalty.

2. Identify the possible learner strategies. This should span all of the possible ways of playing the game to ensure a more complete understanding of the reward signal. However, there may be times when only a small subset of strategies is appropriate to analyze. In general, all possible strategies should be explored when the intended learner is a novice. Whereas, the training developer may limit the scope for analysis if the intended learner is an expert and will focus their decisions on a smaller decision space. Additionally, if the training objectives call for a specific action to take place at a specific time or event in the scenario, this can also be evaluated.

3. Identify which of the possible learner strategies should be rewarded and which strategies should be penalized. This does not have to be precise at this point, but can assist with identifying what possible learner strategies should be evaluated. This analysis should explicitly reflect the training objectives.

4. Develop the means to batch run the games with an automated tool. This may result in considerable amount of work if it is not created already. Ideally, the game should be able to run automatically from the command line.

5. Run the game and collect the data. The data collected should identify the strategy or policy used and the result. The result may be a score, a

quantifiable outcome, and any other means of quantifying performance. The result used should mirror the result that the learner will see as a part of the game's performance feedback mechanism. Using the brute force method, a minimum of 30 runs of each strategy is desirable to use the central limit theorem (CLT) as a part of the analysis. Using a reinforcement-learning approach requires some iterative experiments to determine how long it takes the reinforcement-learning algorithm to learn the environment and determine higher rewarded strategies and policies.

6. Analyze the data. Use a statistical analysis software package to understand the mean and standard error of each strategy. Organize the results in rank order. Then compare the different strategies to each other. Look at the list of strategies and determine if 1) only acceptable strategies are among the highest rewarded strategies and 2) only unacceptable strategies are among the least rewarded strategies. This ensures that good performance is rewarded and poor performance is penalized.

7. Adjust the scenario or reward function of the game or scenario as needed. If bad performance is inadvertently rewarded or good performance is penalized, there is a problem with the scenario or game that produces this result. The scenario designer must redo the experimental runs after any changes are made to the scenario or game to ensure no inadvertent mistakes were made during the editing.

B. TECHNICAL APPROACH

The UrbanSim game is composed of the graphical user interface that is unique to UrbanSim. Within the UrbanSim game, PsychSim is the simulation model that is used to adjudicate the user actions and impact on the game environment. Python code from David Pynadath, was modified to interface with the UrbanSim's PsychSim software to conduct the experiments. This code enabled the simulation experiments to run from the command line, which in turn enabled batch running as well as reducing the time to play the game from roughly an hour per game to approximately one minute. Figure 25 describes the existing UrbanSim practice environment and the software components added to execute the experiments.

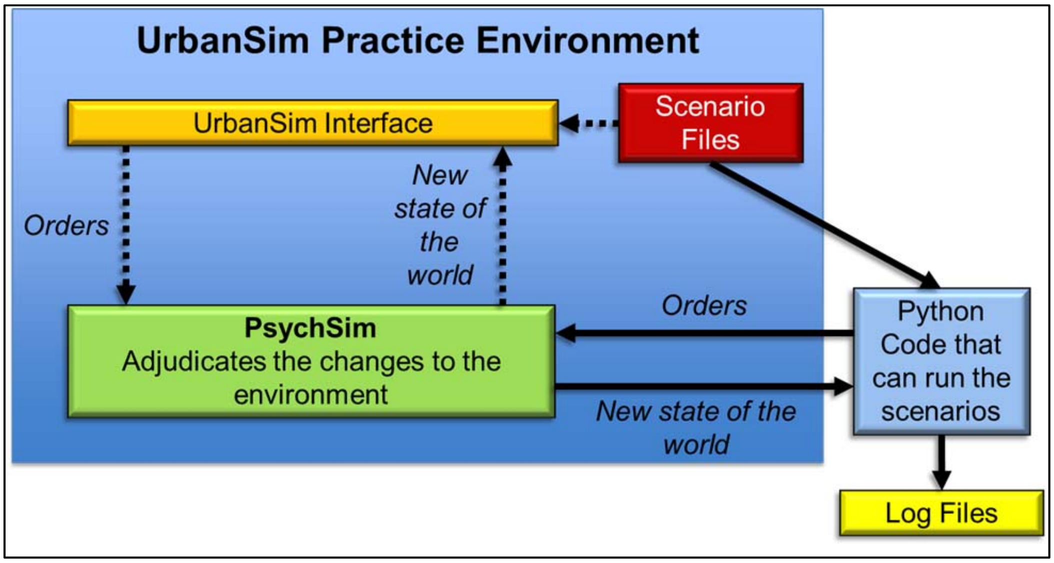

Figure 25. The experiment configuration

C. THREE-DIGIT STRATEGY CODE BATCH EXPERIMENT

The first iteration of the test focused on a simple strategy approach. One of the education objectives of UrbanSim is to reinforce the ‘‘Clear, Hold, Build’’ approach to counterinsurgency, as outlined in FM 3-24, Counterinsurgency Operations. The PsychSim software uses a library function that contains the ‘‘object,’’ the ‘‘type,’’ and the ‘‘actor.’’ The ‘‘object refers to the area, structure, unit, or individual that is acted upon, such as ‘‘Kassad Quarter,’’ ‘‘Shipping Terminal,’’ ‘‘Tribe 1,’’ or ‘‘Asad.’’ The ‘‘type’’ refers to the verb of action that will occur, such as ‘‘Arrest Person,’’ ‘‘Repair,’’ or ‘‘Patrol Neighborhood.” The ‘‘actor’’ refers to the agent that will do the ‘‘type’’ to the ‘‘object,’’ such as ‘‘H Co A,’’ ‘‘Battalion Commander,’’ or ‘‘CA Unit.’’ Using this library, each agent’s available actions were binned in one of three bins. The three bins contain actions that are associated with clear, hold, and build. Each possible action was put in a bin by evaluating the agent and sorting them by ‘‘type.’’ The ‘‘type’’ in each agent’s action list refers to a verb such as ‘‘cordon and search’’ or ‘‘host meeting.’’ The following chart describes where all of the actions were binned.

Table 2. List of verbs used to bin available actions as Clear, Hold and Build. Note that ‘‘Give Propaganda’’ is used in PsychSim but this action is called ‘‘Information Engagement’’ in UrbanSim

Clear	Hold	Build
Seize Structure	Joint Investigate	Repair
Cordon and Knock	Recruit Soldiers	Recruit Soldiers
Cordon and Search	Recruit Police	Recruit Police
Dispatch Individual	Advise	Advise
Attack Group	Set up Checkpoint	Arrest Person
Set up Checkpoint	Remove	Give Gift
Remove	Arrest Person	Host Meeting
Arrest Person	Give Gift	Support Politically
Give Gift	Host Meeting	Pay
Host Meeting	Support Politically	Treat Wounds/Illness
Support Politically	Pay	Patrol Neighborhood
Patrol Neighborhood	Treat Wounds/Illness	Give Propaganda*
Give Propaganda*	Patrol Neighborhood	
	Give Propaganda*	

From these bins 27 different strategies were developed which represent the 27 possible combinations of ‘‘c,’’ ‘‘h,’’ and ‘‘b.” The strategy consists of an approach for the first five turns, the second five turns, and the last five turns. For each game, the agent was given one of the 27 generated strategies, such as ‘‘chb’’ which represents clear tasks for the first five turns, hold tasks for the middle five turns, and build tasks for the final five turns. No other selection criteria was used to determine the player's actions outside of the 27 derived courses of action. Each of the 27 approaches was replicated 37 times.

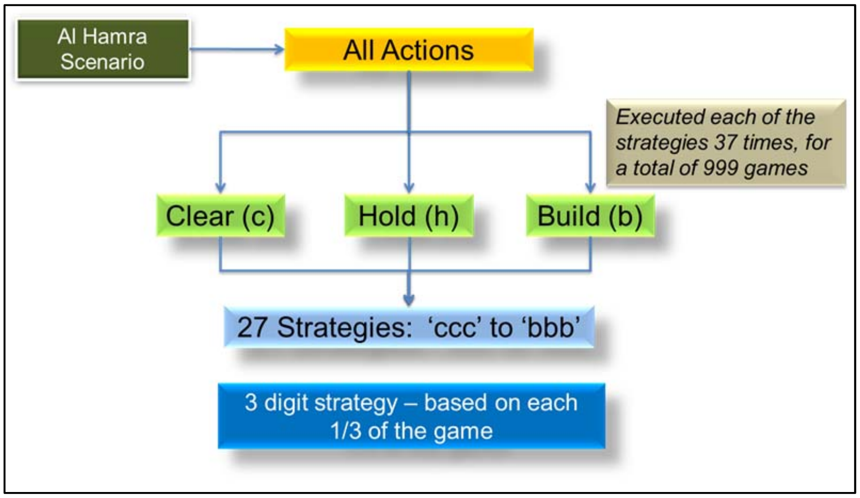

Figure 26. 3-Digit strategy development

D. FIVE-DIGIT STRATEGY CODE BATCH EXPERIMENT

The 3-digit experiment provided insight concerning the ''Clear, Hold, Build'' training objective. However, the 3-digit experiment did not provide any insights about the ''lethal versus non-lethal versus mixed lethal and non-lethal'' training objective or the ''legal versus illegal'' training objectives. Therefore, a 5-digit strategy code was developed and tested.

After analyzing results of the 3-digit experiment, it appeared that ''Clear'' tasks were penalized more than expected. A closer analysis of the tasks associated with each bin revealed that many of the actions in the Clear bin were actions that could be considered violations of the Law of Land Warfare. For example, ''dispatching'' (killing) the mayor, removing the hospital, attacking a region, and seizing the city''s municipal building were in this bin. Further analysis revealed that 47% of the clear actions were illegal in nature, whereas, 29% of the hold actions and 34% of the build actions were illegal in nature.

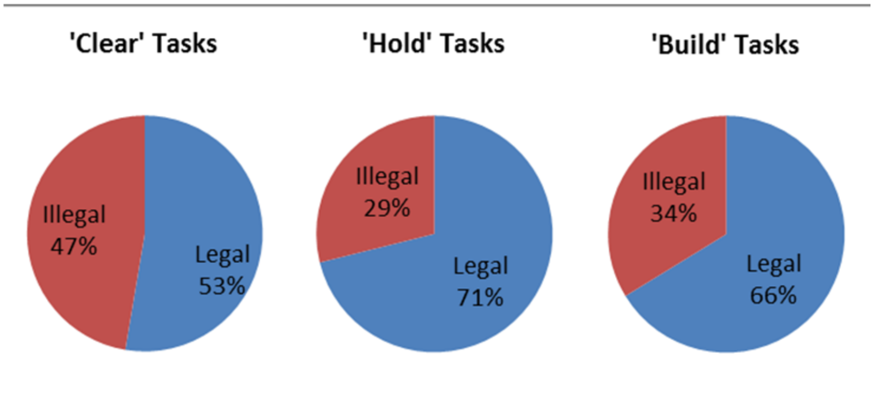

Figure 27. Pie chart of ‘‘Clear,’’ ‘‘Hold, ‘‘Build’’ Tasks

Therefore, the strategies where further binned by exclusively legal and all available actions (which included all legal and illegal actions).

The next iteration of the experiment sought to address the next two research questions: 1) Does the scenario reward student actions that are exclusively legal over student actions that are mixture of legal and illegal actions? and 2) Does the scenario reward student actions that are a mixture of lethal and non-lethal actions over exclusively lethal or exclusively non-lethal?

To address these questions, a 5-digit strategy code was developed.

The first digit determined if the strategy was exclusively ‘‘legal’’ or included both ‘‘legal’’ and ‘‘illegal’’ actions. An analysis of the scenario file enabled categorizing the actions as ‘‘legal’’ and ‘‘illegal.’’ For purposes of this experiment it was decided that ‘‘killing’’ a friendly actor is ‘‘illegal’’ but ‘‘killing’’ a bad actor is ‘‘legal.’’ To discern these differences, a list of ‘‘opposing actors/facilities’’ were determined from the scenario file. ‘‘Opposing actors/facilities’’ were defined as things, people, or groups that opposed coalition efforts. Table 2 lists the Opposing Actors/Facilities with the associated reason for this determination.

Table 3. List of Opposing Actors/Facilities

Opposing Actors/Facilities	Reasoning
Asad	Enemy Sniper
Firing Range	Population does not support it
Granary 2 (IED Manufacturing Plant)	Produces IEDs
Weapons Cache (business)	Weapons Cache
Al-Qassas Brigade Safehouse	Supports the enemy Al Qassas Brigade
JAAS Safehouse	Supports the enemy JAAS
Kurdish Raiders	Opposes HN and Coalition forces
Shiite Death Squads	Oppose HN and Coalition forces
Weapons Cache (home)	Weapons Cache
Shiite Death Squad Safehouse	Supports the Shiite Death Squads
JAAS	Opposes HN and Coalition forces
Al-Qassas Brigade	Opposes HN and Coalition forces

The next step determined if the action was positive or negative in nature. Table 3 lists the actions that were assessed to be positive or negative in nature. Illegal actions were defined as ‘‘positive’’ actions for ‘‘opposing actors/facilities’’ and ‘‘negative’’ actions for non-’’opposing actors/facilities.’’ Legal actions were defined as ‘‘negative’’ actions for ‘‘opposing actors/facilities’’ and ‘‘positive’’ actions for non-’’opposing actors/facilities.’’

Table 4. List of negative and positive actions

Negative	Positive
Arrest Person	Advise
Attack Group	Cordon and Knock
Dispatch Individual	Cordon and Search
Remove	Give Gift
Seize Structure	Host Meeting
	Information Engagement
	Joint Investigate
	Patrol Neighborhood
	Pay
	Recruit Police
	Recruit Soldiers
	Release Person
	Repair
	Set up Checkpoint
	Support Politically
	Treat Wounds/Illnesses

The next digit determined if the strategy was ‘‘lethal’’ or ‘‘nonlethal,’’ or a mix of ‘‘lethal’’ and ‘‘non-lethal.’’ It is a subjective assessment if an action was determined ‘‘lethal’’ or ‘‘nonlethal.’’ Table 4 lists the type of actions that are ‘‘lethal’’ and ‘‘nonlethal.’’ For some actions, such as ‘‘arrest person,’’ it was subjectively determined that this is a lethal action because it removed that entity from the environment.

Table 5. Actions that are Lethal and Nonlethal

Lethal	Nonlethal
Arrest Person	Advise
Attack Group	Give Gift
Cordon and Knock	Give Propaganda
Cordon and Search	Host Meeting
Dispatch Individual	Information Engagement
Joint Investigate	Pay
Patrol Neighborhood	Recruit Police
Remove	Recruit Soldiers
Seize Structure	Release Person
Set up Checkpoint	Repair
	Support Politically
	Treat Wounds/Illnesses

The last three digits were the same as the 3-digit strategy code; ‘‘Clear,’’ ‘‘Hold,’’ or ‘‘Build’’ for the first, middle, and last five turns of the 15 turn game. There are 162 distinctly different strategies associated with the 5-digit strategy code. Each of the 162 strategies was executed 30 times for a total of 4,860 games.

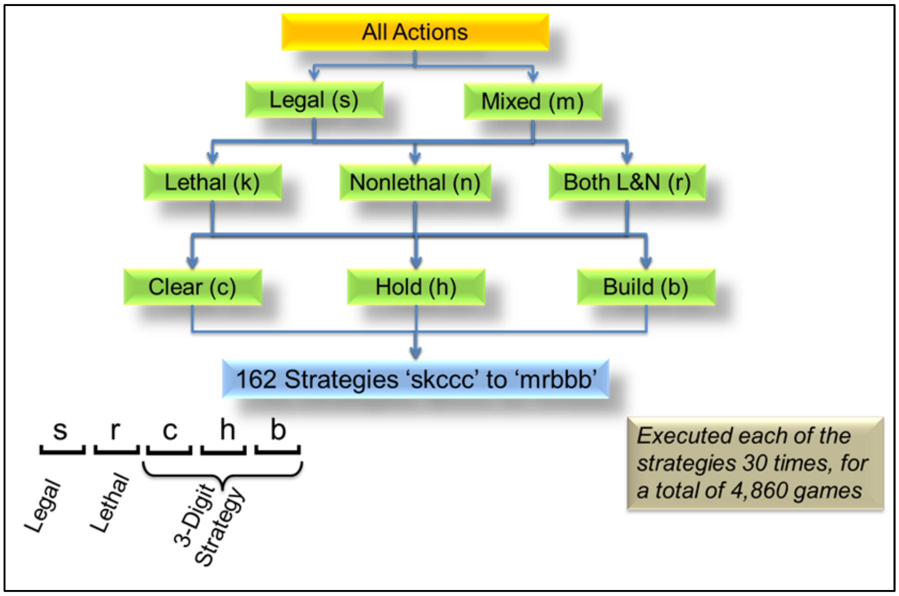

Figure 28. 5-digit strategy development.

E. FIVE-DIGIT STRATEGY CODE REINFORCEMENT-LEARNING EXPERIMENT

This experiment used the same 162 different strategies that were used in the 5-digit batch experiment. However, instead of running 30 iterations of each strategy, a reinforcement-learning algorithm explored and gained insight about the underlying reward structure. The experiment used an epsilon-greedy strategy for the exploratory policy. The epsilon-greedy strategy selects the best strategy with a proportion of 1- of the number of trials. The value for   was 0.1, which determines that 10% of the time, the agent will take a randomly selected strategy, and 90% of the time the agent will select the highest valued strategy. The experiment used the Direct-Q Computation (DQ-C) method for the value function. The reward function was the end of 15-turn game.

The experiment ran for 10,000 iterations with the first 5,000 iterations using a randomly selected policy. The last 5,000 iterations used an increasingly greedy strategy selection. The key data collected from this experiment is the value estimates of the strategies. The value estimate of the strategy is the discounted average of the scores of the previous games using the particular strategy. The value estimate is not the expected score of the strategy.

This experiment provides unique insight about the reward structure that is not evident from the batch runs. The reinforcement-learning experiment provides the scenario designer information about the strength of the reward signal compared to the noise. This experiment seeks to determine if the reward signal is strong enough for the learner to differentiate between optimal and non-optimal strategies.

IV. RESULTS AND DISCUSSION

A. DOES URBANSIM'S PERFORMANCE FEEDBACK SYSTEM SUPPORT THE STATED LEARNING OBJECTIVES?

1. Does the Al Hamra Scenario Reward the ''Clear, Hold, Build'' Approach Over the Other Approaches?

The following chart depicts the distribution of outcomes from the 3-digit batch experiment. From this plot, the highest rewarded 3-digit strategy is ''bbb,'' which represents ''build, build, build'' and the most penalized strategy is ''ccc,'' which represents ''clear, clear, clear'' for each third of the game. Figure 30 is a plot of the strategy's mean score with standard error bars. From these outcomes, a Tukey-Kramer HSD analysis of the data shows which strategy scores are significantly different from other strategies.

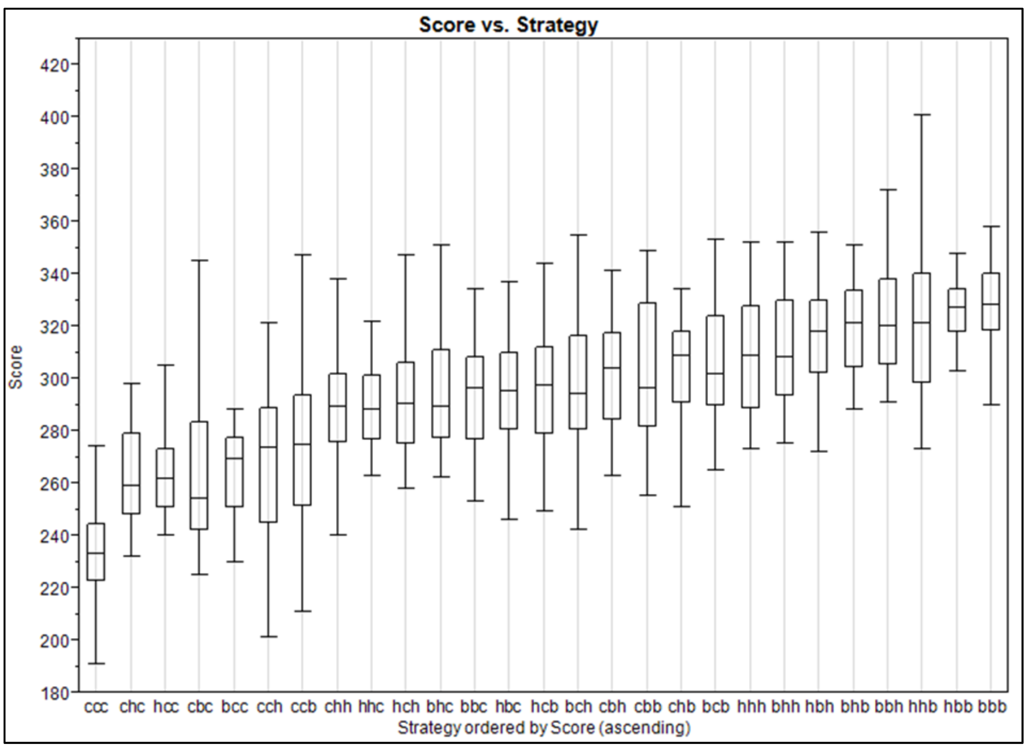

Figure 29. Boxplot of the results of the 3-Digit Strategy Experiment.

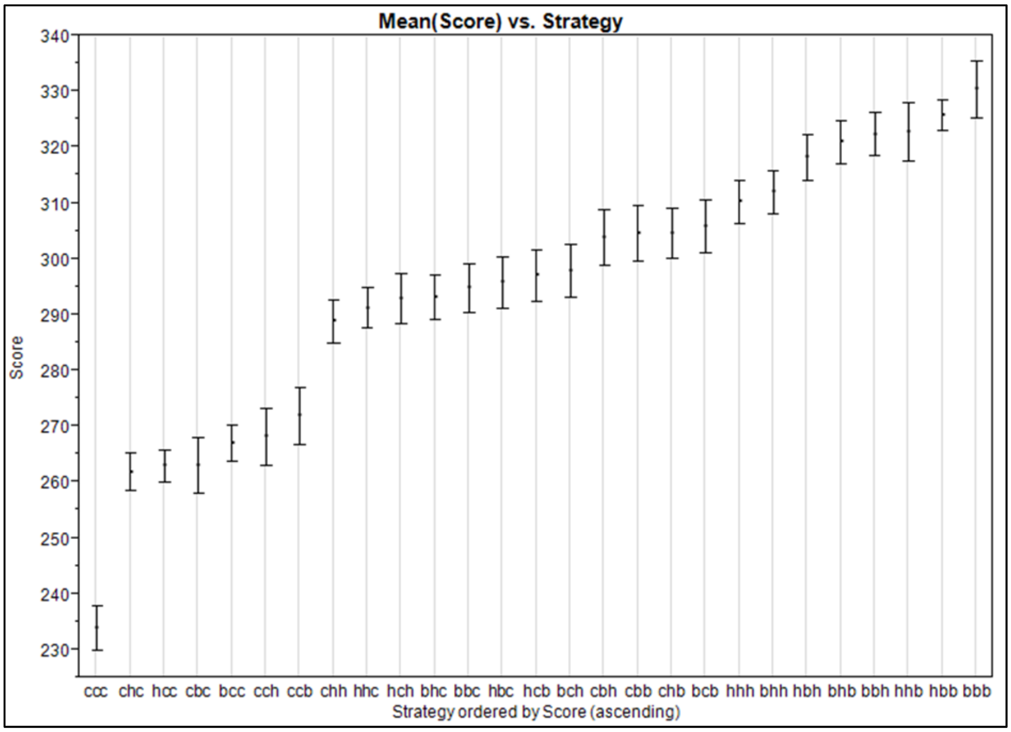

Figure 30. **Plot of the Mean Score vs Strategy with standard error bars.**

Table 6. Tukey-Kramer HSD Connecting Letters Report that depicts which strategies are significantly different from each other.

Connecting Letters Report

Level		Mean
bbb	A	330.50000
hbb	A B	325.80000
hhb	A B	322.80000
bbh	A B	322.36667
bhb	A B	320.96667
hbh	A B C	318.33333
bhh	A B C D	312.00000
hhh	A B C D E	310.30000
bcb	B C D E	305.96667
chb	B C D E	304.66667
cbb	B C D E	304.56667
cbh	B C D E	303.83333
bch	C D E	297.86667
hcb	C D E	297.06667
hbc	C D E	295.86667
bbc	D E	294.83333
bhc	D E F	293.30000
hch	D E F	293.00000
hhc	D E F	291.26667
chh	E F G	288.86667
ccb	F G H	271.96667
cch	G H	268.20000
bcc	G H	267.13333
cbc	H	263.03333
hcc	H	262.96667
chc	H	261.90000
ccc	I	233.96667

Levels not connected by same letter are significantly different.

From this data, the scenario designer would assess the results comparing them to the desired training objectives. First, the scenario designer would look at the highest rewarded strategies that are similar, determine if they contain acceptable strategies and do not contain unacceptable strategies. Second, look at the least rewarded strategies, determine if they contain unacceptable strategies and do not contain acceptable strategies. This tests the results using the method depicted in Figures 6 and 7.

2. Does the Scenario Reward Student Actions that are Exclusively Legal Over Student Actions that are a Mixture of Legal and Illegal Actions?

Figure 31 is a plot that depicts the outcome of the 162 5-Digit strategies.

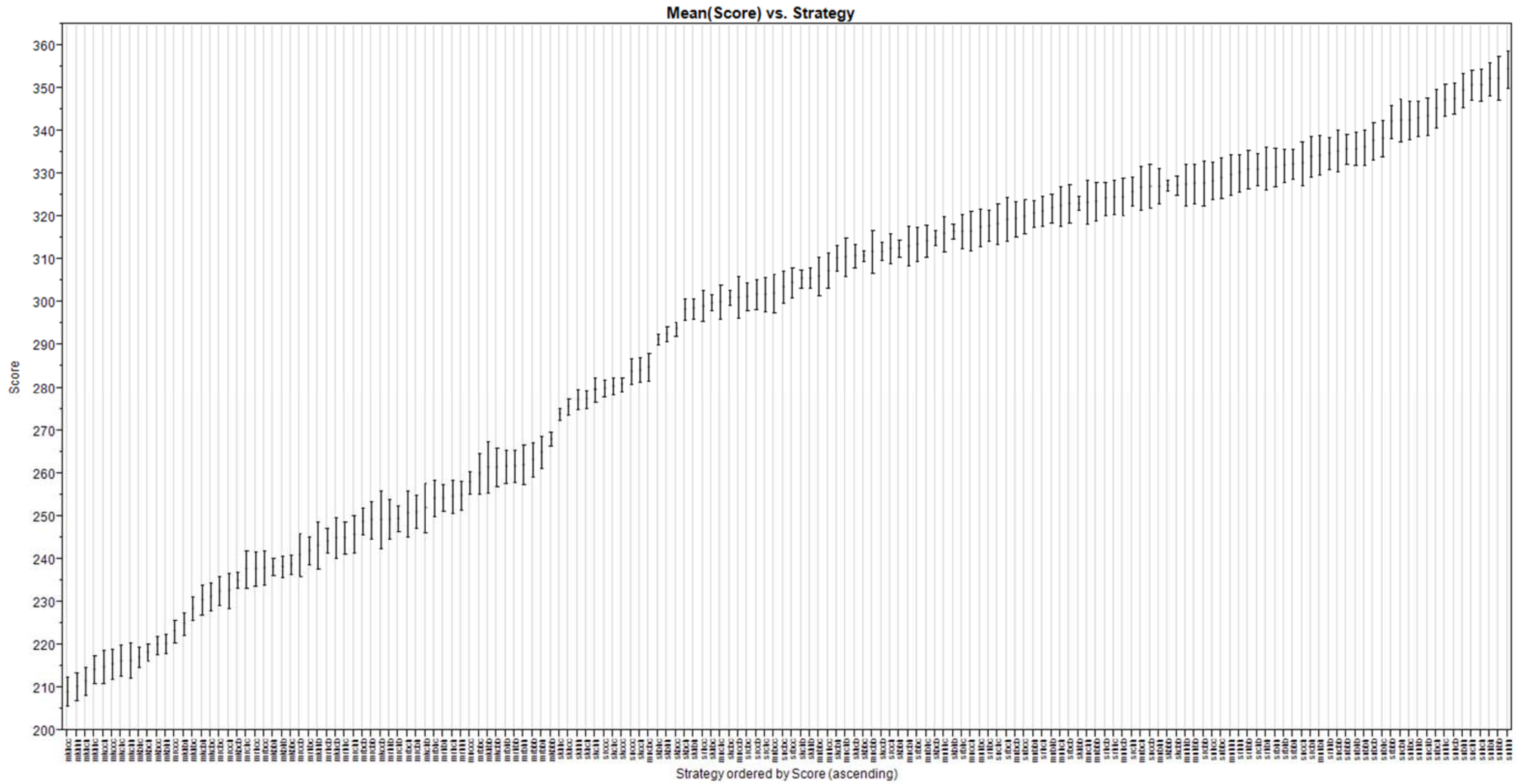

Figure 31. Plot mean and standard error bars of the 5-Digit strategies.

Tables 3–6, 5-digit Strategy results, lists the best scoring strategy, the mean, and the other strategies that are not significantly different (denoted by a darkened vertical block with a common heading number).

Table 7. Five-digit Strategy results, strategies 1 – 45. Strategies that share a common shaded block, by number, are not rewarded significantly different

	Strategy Code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	Mean Score
1	snhhh	■																										354.40
2	snhbb	■	■																									352.37
3	snhbh	■	■																									352.20
4	snhch	■	■																									350.83
5	snchh	■	■																									350.77
6	snbhh	■	■																									349.50
7	snhcb	■	■	■																								347.63
8	snhhc	■	■	■																								347.33
9	snbch	■	■	■	■																							345.33
10	snchb		■	■	■	■																						343.47
11	snhhb		■	■	■	■	■																					342.97
12	snhbc		■	■	■	■	■	■																				342.63
13	sncbh		■	■	■	■	■	■	■																			342.50
14	srbbb		■	■	■	■	■	■	■	■																		342.17
15	snbhc			■	■	■	■	■	■	■	■																	338.33
16	snbcb			■	■	■	■	■	■	■	■	■																337.67
17	snbbh				■	■	■	■	■	■	■	■	■															336.20
18	snbhb				■	■	■	■	■	■	■	■	■															335.90
19	snbbb				■	■	■	■	■	■	■	■	■															335.77
20	sncbb				■	■	■	■	■	■	■	■	■															335.40
21	srhhb					■	■	■	■	■	■	■	■	■														334.73
22	mnhbh					■	■	■	■	■	■	■	■	■														334.33
23	srcbh					■	■	■	■	■	■	■	■	■	■													333.97
24	sncch						■	■	■	■	■	■	■	■	■	■												332.47
25	srbbh							■	■	■	■	■	■	■	■	■												332.30
26	srbhb								■	■	■	■	■	■	■	■												331.97
27	srbhh									■	■	■	■	■	■	■	■											331.63
28	srhbh										■	■	■	■	■	■	■											331.33
29	srchb										■	■	■	■	■	■	■	■										331.13
30	srhbb										■	■	■	■	■	■	■	■										331.07
31	srhhh										■	■	■	■	■	■	■	■	■									330.23
32	mnhhh										■	■	■	■	■	■	■	■	■	■								329.70
33	snbbc										■	■	■	■	■	■	■	■	■	■								329.03
34	snhcc										■	■	■	■	■	■	■	■	■	■	■							328.40
35	srcbb										■	■	■	■	■	■	■	■	■	■	■	■						327.87
36	mnhbb											■	■	■	■	■	■	■	■	■	■	■						327.70
37	mnhhb											■	■	■	■	■	■	■	■	■	■	■						327.43
38	skbbb											■	■	■	■	■	■	■	■	■	■	■						327.27
39	skcbb											■	■	■	■	■	■	■	■	■	■	■						327.27
40	mnbhh											■	■	■	■	■	■	■	■	■	■	■	■					327.17
41	snccb											■	■	■	■	■	■	■	■	■	■	■	■	■				327.13
42	mnbch												■	■	■	■	■	■	■	■	■	■	■	■				326.77
43	srchh												■	■	■	■	■	■	■	■	■	■	■	■	■			325.83
44	mnhcb													■	■	■	■	■	■	■	■	■	■	■	■	■		324.60
45	srhhc													■	■	■	■	■	■	■	■	■	■	■	■	■		324.50

Table 8. Five-digit Strategy results, strategies 46 – 90. Strategies that share a common shaded block, by number, are not rewarded significantly different.

	Strategy Code	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	Mean Score
46	srhcb		■	■	■	■	■	■	■	■	■	■	■	■	■																					324.23
47	mnbbb			■	■	■	■	■	■	■	■	■	■	■	■	■																				323.50
48	mnchh			■	■	■	■	■	■	■	■	■	■	■	■	■																				323.43
49	skhbb				■	■	■	■	■	■	■	■	■	■	■	■	■																			323.10
50	srbcb				■	■	■	■	■	■	■	■	■	■	■	■	■																			323.07
51	mnhch				■	■	■	■	■	■	■	■	■	■	■	■	■																			322.47
52	mnbhb				■	■	■	■	■	■	■	■	■	■	■	■	■	■																		321.97
53	srhch					■	■	■	■	■	■	■	■	■	■	■	■	■	■																	321.33
54	mnbbh						■	■	■	■	■	■	■	■	■	■	■	■	■	■																320.73
55	snbcc							■	■	■	■	■	■	■	■	■	■	■	■	■																320.10
56	mnbcb								■	■	■	■	■	■	■	■	■	■	■	■																319.50
57	srbch								■	■	■	■	■	■	■	■	■	■	■	■																319.43
58	snchc									■	■	■	■	■	■	■	■	■	■	■																318.33
59	srhbc									■	■	■	■	■	■	■	■	■	■	■	■															317.90
60	mnhbc										■	■	■	■	■	■	■	■	■	■	■															317.47
61	mncch											■	■	■	■	■	■	■	■	■	■															316.67
62	srbhc												■	■	■	■	■	■	■	■	■															316.60
63	skbhb												■	■	■	■	■	■	■	■	■															316.60
64	mnhhc													■	■	■	■	■	■	■	■	■														316.00
65	skbcb														■	■	■	■	■	■	■	■	■													315.07
66	mnbhc														■	■	■	■	■	■	■	■	■													314.33
67	srbbc															■	■	■	■	■	■	■	■	■												313.67
68	mncbh															■	■	■	■	■	■	■	■	■												313.10
69	skbbh																■	■	■	■	■	■	■	■	■											312.60
70	srcch																■	■	■	■	■	■	■	■	■											312.57
71	skccb																	■	■	■	■	■	■	■	■	■										311.87
72	mncbb																	■	■	■	■	■	■	■	■	■										311.87
73	skbbc																		■	■	■	■	■	■	■	■	■									310.83
74	skhcb																		■	■	■	■	■	■	■	■	■									310.80
75	mnchb																			■	■	■	■	■	■	■	■	■								310.57
76	skcbh																			■	■	■	■	■	■	■	■	■	■							310.37
77	mnhcc																				■	■	■	■	■	■	■	■	■	■						307.43
78	mnbbc																					■	■	■	■	■	■	■	■	■						306.03
79	skhhb																					■	■	■	■	■	■	■	■	■						305.70
80	skchb																					■	■	■	■	■	■	■	■	■						305.53
81	srbcc																						■	■	■	■	■	■	■	■						304.57
82	sncbc																							■	■	■	■	■	■	■	■					303.57
83	mnbcc																								■	■	■	■	■	■	■	■				302.17
84	srchc																									■	■	■	■	■	■	■	■			301.87
85	srccb																									■	■	■	■	■	■	■	■			301.83
86	srcbc																									■	■	■	■	■	■	■	■			301.40
87	mnccb																										■	■	■	■	■	■	■			301.20
88	skcbc																										■	■	■	■	■	■	■			301.10
89	mnchc																											■	■	■	■	■	■			300.07
90	skhbc																												■	■	■	■	■			299.97

Table 9. Five-digit Strategy results, strategies 91–135. Strategies that share a common shaded block, by number, are not rewarded significantly different.

	Strategy Code	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	Mean Score
91	srhcc																													299.20
92	skhbh																													298.53
93	skbch																													298.37
94	skbcc																													293.77
95	skbhh																													292.60
96	skbhc																													291.37
97	mncbc																													284.80
98	skcch																													284.20
99	snccc																													283.80
100	skccc																													280.79
101	skchc																													280.37
102	srccc																													279.83
103	skchh																													279.57
104	skhch																													277.27
105	skhhh																													277.17
106	skhcc																													275.57
107	skhhc																													273.77
108	mkbbb																													267.97
109	mrbbh																													264.87
110	mrbbb																													263.07
111	mrbhh																													262.00
112	mrhbb																													261.73
113	mrbhb																													261.60
114	mkcbb																													261.47
115	mkhbb																													261.30
116	mrbbc																													259.93
117	mnccc																													257.80
118	mrhhh																													254.80
119	mrhch																													254.57
120	mrhbh																													254.27
121	mrbhc																													254.23
122	mkchb																													251.90
123	mrcbh																													250.97
124	mrbch																													250.57
125	mrchb																													249.50
126	mrhhb																													249.20
127	mkccb																													249.17
128	mrcbb																													249.07
129	mrbcb																													248.80
130	mrchh																													245.77
131	mkhcb																													244.93
132	mrhhc																													244.93
133	mrhcb																													244.23
134	mkhhb																													243.13
135	mrhbc																													241.90

Table 10. Five-digit Strategy results, strategies 136 – 162. Strategies that share a common shaded block, by number, are not rewarded significantly different.

	Strategy Code	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	Mean Score
136	mrccb			■	■	■	■	■	■	■								240.83
137	mkbbc				■	■	■	■	■	■	■							238.70
138	mkbhb					■	■	■	■	■	■							238.23
139	mkbbh					■	■	■	■	■	■							238.20
140	mrbcc					■	■	■	■	■	■							237.93
141	mrhcc					■	■	■	■	■	■							237.73
142	mrchc					■	■	■	■	■	■							237.60
143	mkbcb						■	■	■	■	■	■						235.03
144	mrcch							■	■	■	■	■	■					232.67
145	mrcbc								■	■	■	■	■					232.57
146	mkcbc								■	■	■	■	■					231.13
147	mkcbh								■	■	■	■	■	■				230.50
148	mkhbc								■		■	■	■	■	■			228.53
149	mkhbh								■			■	■	■	■	■		224.90
150	mrccc								■				■	■	■	■	■	223.10
151	mkbhh								■					■	■	■	■	220.14
152	mkbcc								■						■	■	■	219.87
153	mkbch								■						■	■	■	218.30
154	mkbhc								■							■	■	217.03
155	mkchh								■							■	■	216.33
156	mkchc								■							■	■	216.33
157	mkccc								■							■	■	215.40
158	mkcch								■							■	■	214.80
159	mkhhc								■								■	214.27
160	mkhch								■									211.40
161	mkhhh								■									210.23
162	mkhcc								■									209.10

From the above plots and charts, the scenario designer would determine if it is acceptable for similarly rewarded strategies given the desired training objectives. This analysis only requires the amount of precision that the scenario developer desires.

To answer the research question of whether the scenario rewards student actions that are exclusively legal over student actions that are mixture of legal and illegal actions, the following boxplot depicts the distribution of strategies between ‘‘Mixed Legal and Illegal’’ and ‘‘Exclusively Legal.”

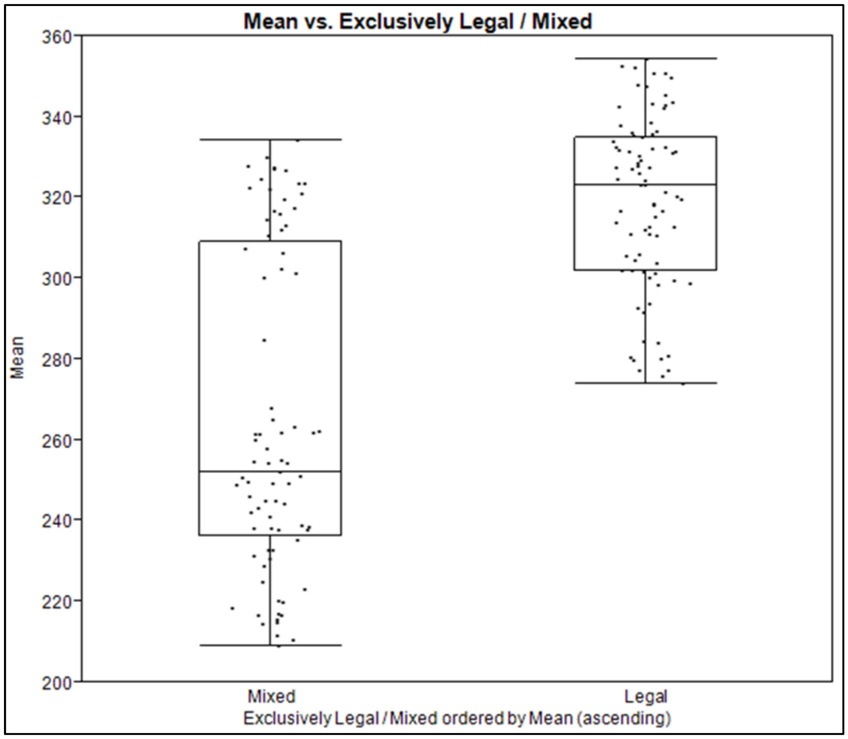

Figure 32. Score vs Exclusively Legal / Mixed Legal and Illegal Actions.

Figure 33 shows the mean of the two groups of strategies and the standard error.

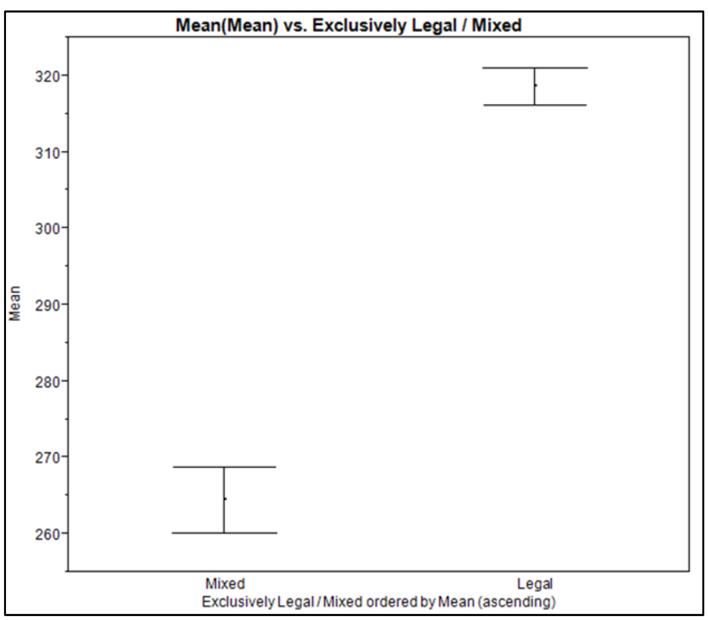

Figure 33. Score vs Exclusively Legal / Mixed Legal and Illegal Actions with mean and standard error bars.

The analysis shows that the strategies that are ‘‘Exclusively Legal’’ are rewarded more than ‘‘Mixed Legal and Illegal.’’

3. Does the Scenario Reward Student Actions that are a Mixture of Lethal and Non-lethal Actions Over Exclusively Lethal or Exclusively Non-lethal?

Figure 34 depicts the distribution of scores of strategies that are ‘‘Lethal,’’ ‘‘Non-lethal,’’ and ‘‘Mixed Lethal and Non-Lethal’’ from the 5-digit strategy experiment.

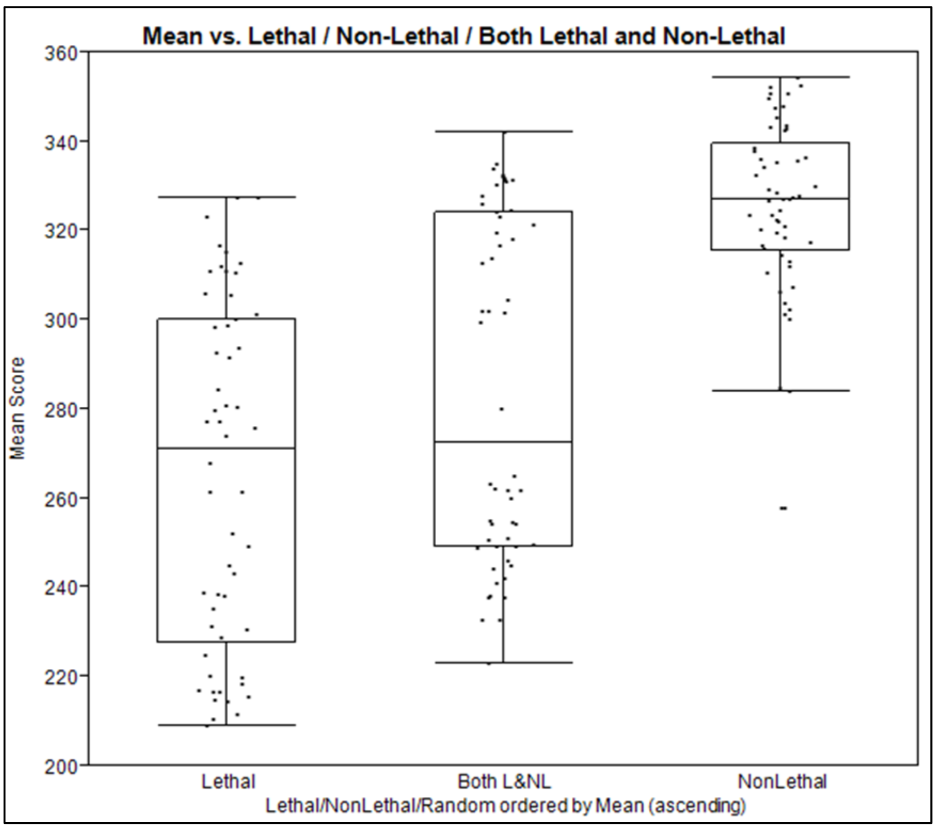

Figure 34. Mean vs. Lethal / Non-Lethal / Both Lethal and Non-Lethal scores box plot.

Figure 35 is a plot of the mean scores associated with the ‘‘Lethal,’’ ‘‘Non-Lethal’’ and ‘‘Mixed Lethal and Non-Lethal’’ strategies.

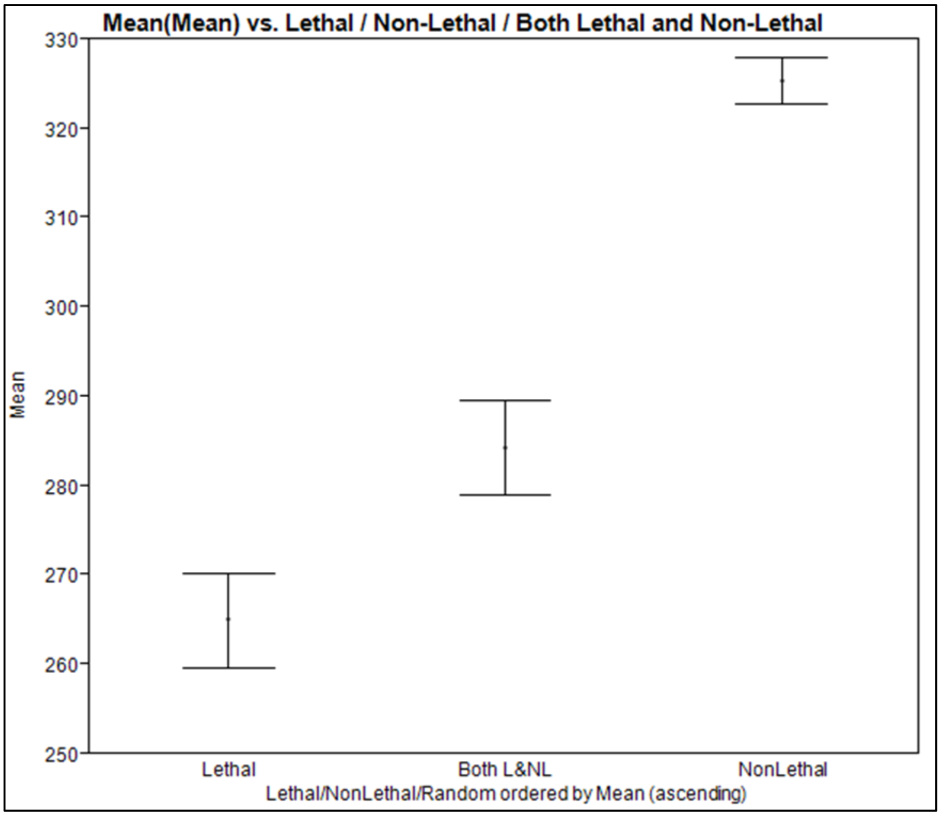

Figure 35. Score vs Lethal, Non-Lethal, and Both Lethal and Non-Lethal actions with standard error bars.

Table 11. Connecting Letters Report from the Lethal, Non-Lethal, and Mixed Lethal and Non-Lethal actions.

Connecting Letters Report

Level				Mean
NonLethal	A			325.42284
Both L&NL		B		284.34568
Lethal			C	264.99431

Levels not connected by same letter are significantly different.

Figures 34, 35, and Table 11 determine that ‘‘Non-Lethal’’ actions are rewarded significantly more than ‘‘Both Lethal and Non-Lethal’’ and ‘‘Lethal,’’ and ‘‘Both Lethal and Non-Lethal’’ is rewarded significantly more than “Lethal.’’ Therefore, if the desired training outcome is to reinforce a mixture of lethal and nonlethal actions, the scenario as written does not adequately reward this policy.

This information would be helpful to the scenario designer or author during the development and creation of the UrbanSim scenario.

4. Is the Performance Feedback Provided to the Learner Strong Enough to Differentiate between Optimal and Non-optimal Strategies?

This research question is addressed using the reinforcement-learning experiment results. The 10,000-iteration experiment estimated the values of the 162 different strategies shown in Table 12.

Table 12. Results of the 162-Strategy Reinforcement-Learning Experiment.

Rank	Strategy	Value
1	s n h b h	345.960
2	m n c b c	341.547
3	s n c b c	335.516
4	s k h c b	330.109
5	m k b c c	324.354
6	m n b h b	316.789
7	s k b c c	313.197
8	m n h h h	308.637
9	s k c b b	307.368
10	m n h h c	307.293
11	s r h h b	306.815
12	m k h h c	306.507
13	s n h c h	304.869
14	m r c c c	304.362
15	s r h h h	303.570
16	m r c h c	299.510
17	s k c b c	298.263
18	s k h c c	298.105
19	s n b h h	298.086
20	m k b h c	297.579
21	m k c h h	297.577
22	m n b c b	297.545
23	s k h b c	297.496
24	s k h h h	297.468
25	m k h c b	297.417
26	m k c c c	297.396
27	s r h h c	297.393
28	s n b b b	297.233
29	m r b h h	297.220
30	s k b h b	297.151
31	s n c h c	297.115
32	s r c h h	297.107
33	s n b c b	297.095
34	m r h b b	296.970
35	s k c c h	296.960
36	s r h c b	296.797
37	m n b b h	296.638
38	m r b c b	296.619
39	m n c c h	296.577
40	m n h h b	296.528
41	s n h c c	296.526
42	m r h h b	296.472
43	m n c h c	296.470
44	m k b b c	296.450
45	m n h b b	296.432
46	m n c h h	296.425
47	s k b b h	296.375
48	s r c b c	296.363
49	s n b c c	296.354
50	s n b h b	296.353
51	s k b b c	296.349
52	s r h c c	296.316
53	m r b h b	296.287
54	m k h b h	296.265
55	s n c h h	296.160
56	s k c b h	296.112
57	s n h h h	296.068
58	m k b h h	296.060
59	m k h h h	295.978
60	s n c c h	295.965
61	m r c c b	295.946
62	s k h b h	295.868
63	m n b b b	295.849
64	m n h c c	295.815
65	m n h c h	295.796
66	s r b b c	295.717
67	m r b h c	295.716
68	s r b b h	295.596
69	m n h b c	295.458
70	m r h c h	295.427
71	s r c c h	295.424
72	s k b h c	295.350
73	s n c c c	295.305
74	m k c c h	295.303
75	m k c b c	295.236
76	s r b h h	295.044
77	m n c c c	295.034
78	s k h b b	294.986
79	s k c h b	294.934
80	m k h b b	294.825
81	s n b b c	294.774
82	s n b c h	294.769
83	s n c b h	294.677
84	m r c h b	294.655
85	m k c h b	294.501
86	m k b b b	294.437
87	s k b c h	294.295
88	m n c b b	294.266
89	m k b b h	294.262
90	s k b c b	294.214
91	m n h c b	294.207
92	m r h c b	294.179
93	s n b h c	294.140
94	m k c b h	294.115
95	s r c b b	294.104
96	s n c h b	294.054
97	s r c h b	293.948
98	m k h b c	293.807
99	m r b c c	293.753
100	m n b h c	293.664
101	s r b h c	293.621
102	m r h b h	293.397
103	s k b h h	293.383
104	m r c b c	293.301
105	m r c h h	293.258
106	s n c b b	293.206
107	s r c h c	293.160
108	m r h h c	293.132
109	m k h h b	293.090
110	m n b b c	293.009
111	m r c c h	292.921
112	m r c b b	292.891
113	s r h b b	292.828
114	m r c b h	292.543
115	m n b c c	292.383
116	s n h h b	292.383
117	s n c c b	292.291
118	m n c b h	292.274
119	m k b c b	292.229
120	m k c b b	292.193
121	m k h c c	292.183
122	m k b c h	292.123
123	s r c c c	291.946
124	m k c c b	291.934
125	m r h b c	291.927
126	s n h b b	291.896
127	s k c c c	291.721
128	m r b b c	291.525
129	s r h b c	291.429
130	s r b h b	291.260
131	s r b c b	291.183
132	s k c c b	291.030
133	m n b c h	290.896
134	m n c c b	290.622
135	m n b h h	290.426
136	s r b b b	290.245
137	s r c c b	290.227
138	s n h b c	288.160
139	s r c b h	288.145
140	s k c h h	287.417
141	s k h h c	286.888
142	s k b b b	286.857
143	m n h b h	286.807
144	m r h c c	286.766
145	s n h c b	286.724
146	s r b c h	286.472
147	s k h c h	286.420
148	m r h h h	286.222
149	m r b c h	285.948
150	m k c h c	285.840
151	m r b b h	285.635
152	m r b b b	285.612
153	s k h h b	285.593
154	m k b h b	284.990
155	s r h b h	284.891
156	s n h h c	284.592
157	s r h c h	284.416
158	s n b b h	283.754
159	m n c h b	283.157
160	s r b c c	283.124
161	s k c h c	283.017
162	m k h c h	282.886

The ranked strategies using the batch method and the reinforcement learning approach are different. This indicates that there is a large ratio of noise to signal for this scenario. The scenario designer can use this information to reduce the noise associated with the reward signal to speed learning for novice students. Conversely, the scenario designer could increase the noise associated with the reward signal to challenge more experienced students.

Figure 36 is a plot of the strategy the reinforcement-learning agent used for each game. The strategy was selected randomly for the first 5,000 games. After the 5,000th game, the selected strategy was increasingly more greedy.

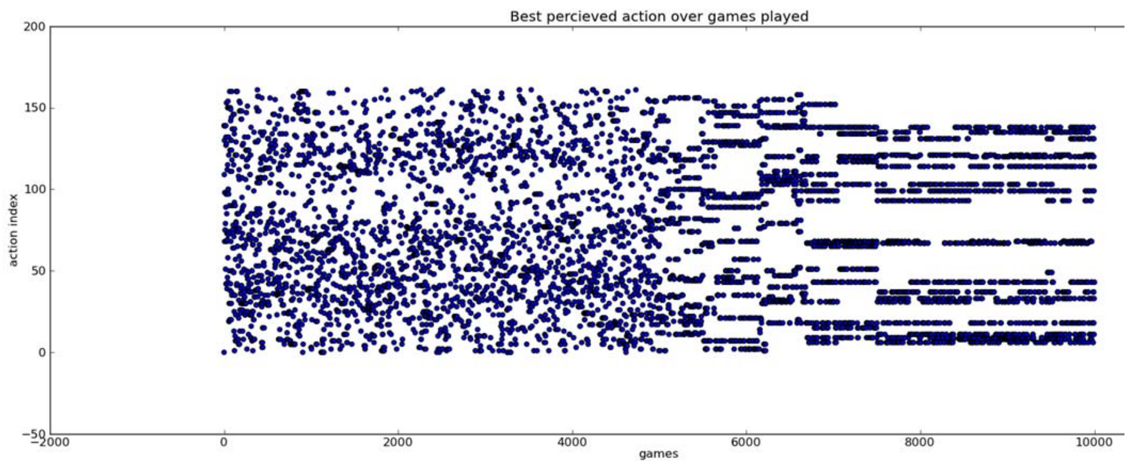

Figure 36. The Best perceived action over the games played. The x-axis is the game number and the y-axis is the strategy index number.

An analysis of the strategies the reinforcement-learning agent valued the most over the number of games played provides some insight about the reward structure. Figures 37 to 41 are histograms of the number of times the reinforcement-learning agent identified a strategy to be the most valuable. The batch run experiments demonstrated that there was no significant difference in the top 9 strategies. Therefore, it is reasonable that the reinforcement-learning agent identified 15 different strategies as the most valuable in the last 50 games played.

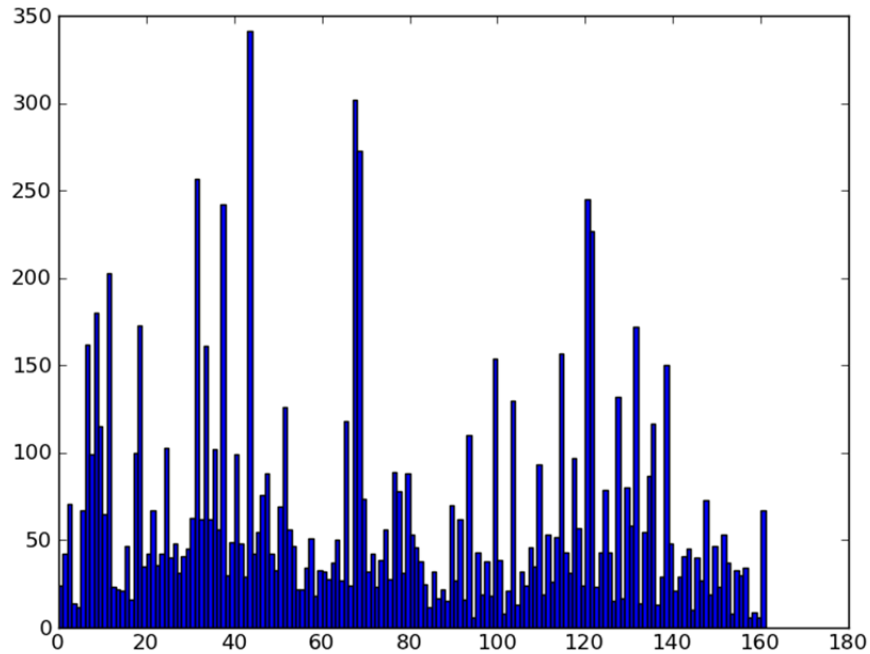

Figure 37. **Histogram of all of the strategies used. The x-axis represents the strategy index number and the y-axis is the frequency the strategy was determined to be the greatest value.**

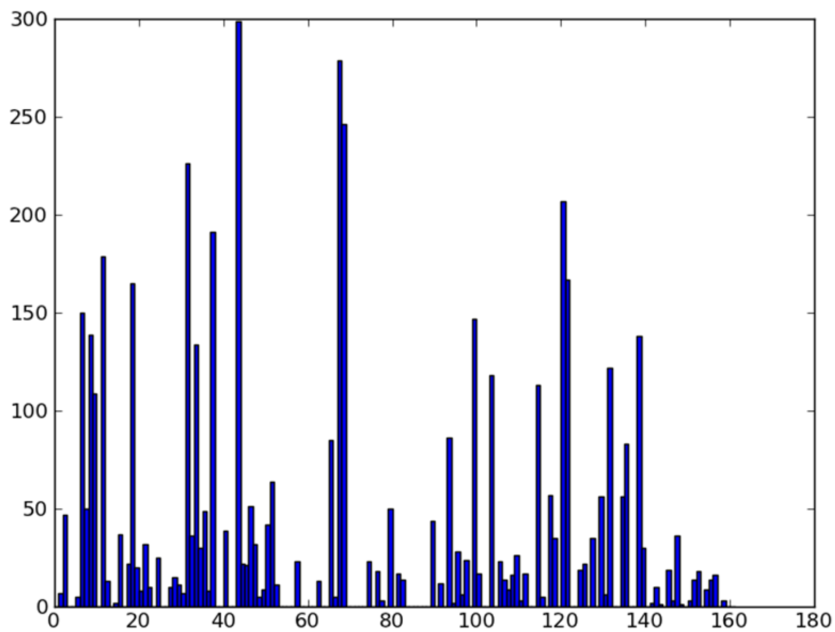

Figure 38. **Histogram of the last 5000 games. The x-axis represents the strategy index number and the y-axis is the frequency the strategy was determined to be the greatest value.**

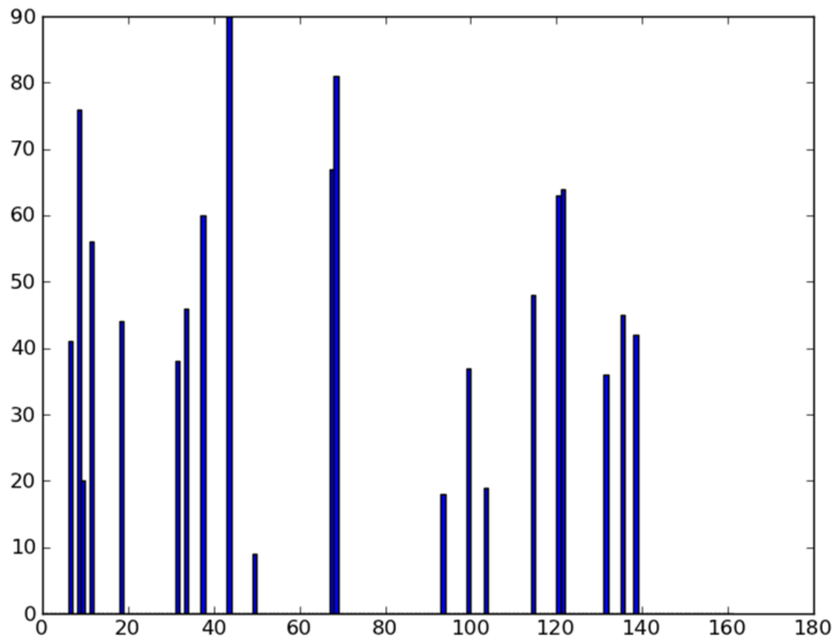

Figure 39. Histogram of the last 1000 games. The x-axis represents the strategy index number and the y-axis is the frequency the strategy was determined to be the greatest value.

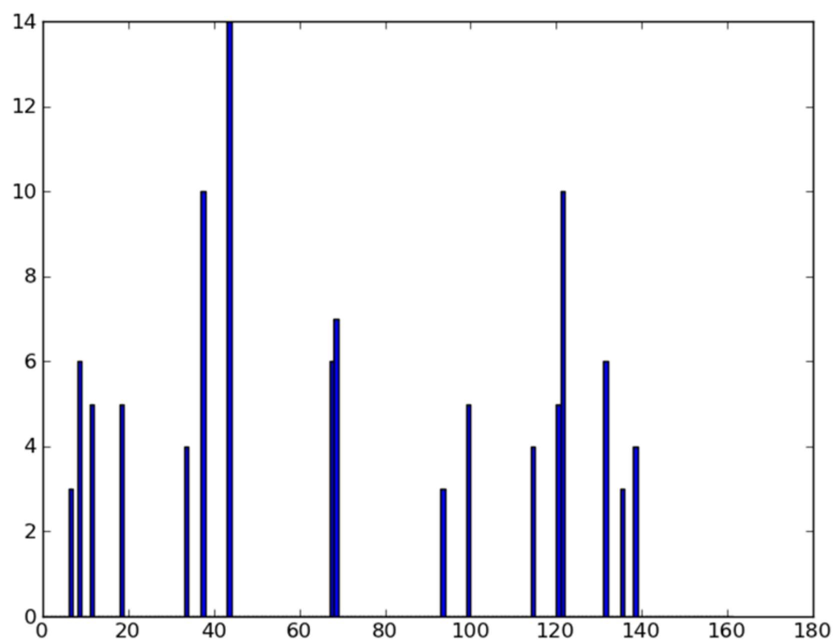

Figure 40. Histogram of the last 100 games. The x-axis represents the strategy index number and the y-axis is the frequency the strategy was determined to be the greatest value.

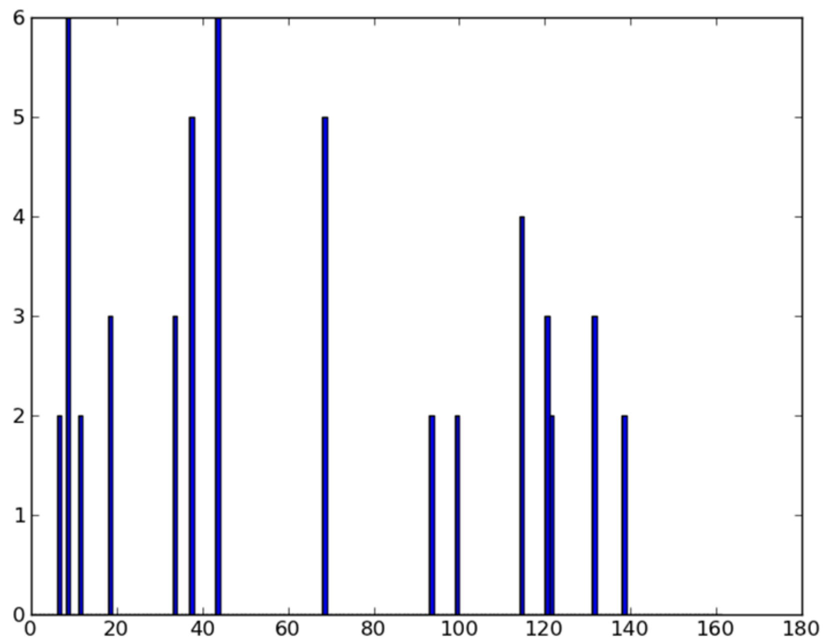

Figure 41. **Histogram of the last 50 games. The x-axis represents the strategy index number and the y-axis is the frequency the strategy was determined to be the greatest value.**

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION AND RECOMMENDATIONS

A. SUMMARY OF RESULTS

This study sought to evaluate the fielded UrbanSim scenarios as they related to the stated training objectives. More generally, this study sought to develop a generalized approach to evaluating scenarios that address ill-defined problems.

From the perspective of evaluating the fielded UrbanSim scenarios, it appears that the unstated, but assumed, training objective of rewarding students that conduct exclusively legal actions is properly rewarded. The training objective of emphasizing the doctrinal principle of ‘‘Clear, Hold, Build’’ did not stand out very clearly. However, it appeared to be in the range of acceptable solutions. The fact that the Build, Build, Build strategy was also in the range of acceptable solutions is not desirable because it reinforces the notion that you can be successful if you ignore the enemy and allow them to operate and you can still be successful in the scenario. The 4th training objective that wants the students to demonstrate that a mixture of lethal and non-lethal actions is better than exclusively lethal or non-lethal was not supported. Non-lethal actions were more strongly rewarded than the mixed approach and the lethal actions. This may be closely tied to the fact that the enemy units in the scenario do not affect the simulated environment enough to replicate the danger of ignoring enemy units operating in the area of operation.

The approach of using automated tools to evaluate a game or game scenario provides insight to the developer and author. Additionally, evaluating a scenario with respect to the training objectives is a necessary step with all training games, but especially true of games that address ill-defined problems. The traditional approach of evaluating scenarios was to define and articulate training objectives, then develop the training scenario, make sure it functions, then use humans to play the scenario, and evaluate the game or scenario based on the training transfer that occurred within the participants. This process is rather resource intensive and can take a considerable amount of time. This approach of using automated tools to evaluate scenarios seeks to reduce the resources and time needed to evaluate training scenarios.

B. GENERALIZABLE RESULTS AND OTHER POTENTIAL APPLICATIONS

In general, this scenario evaluation methodology is able to provide insights about the performance feedback mechanisms in training scenarios that were not available before. The methodology can assist scenario authors throughout the scenario design effort. Similar in nature to the computer programming axiom of ‘‘build a little, test a little,’’ this methodology allows scenario authors to conduct formative, automated testing to ensure the performance feedback mechanism supports the desired training objectives. This methodology provides a means of thoroughly testing and tuning a scenario before human participants begin play testing.

In a different application, this methodology could be applied to evaluating training and education scenarios that address major combat operations. This was the original endeavor of this study, however, it seemed that the decision space was far too large and a game with 15 discrete turns was more manageable. As discussed earlier, the decision space within UrbanSim is deceptively large. Eleven units with between 140 and 341 possible actions over 15 turns generates more than 5×10^{27} possible ways of playing the game. In retrospect, a major combat operations game scenario may be easier to evaluate and provide performance feedback. For a division level scenario there may be 20–25 battalion sized units or units directly controlled by the division which is more than the number of units in UrbanSim. There also may be a few more decision points in the game when the player would give orders. However, for each unit there would be significantly fewer than 341 available actions for each unit, which would drive the decision space down to a manageable level. Using a similar approach of binning actions, the player could give orders to units like ‘‘move’’ to a preidentified location, ‘‘attack’’ an enemy unit, ‘‘shoot indirect fire’’ at an enemy unit, etc., without having to get into the near infinite possibilities of where the unit is moving. Scoping this decision space would not negatively influence the student’s decisions, but would certainly make validating the scenario and providing feedback to the student more manageable.

This methodology also has some potential shortcomings as well. The methodology requires an ability to bin all of the actions available to the learner. For example, the 3- and 5-digit strategy experiments, as well as the reinforcement-learning approach experiments required an ability to bin potential learner actions in Clear, Hold, and Build bins in addition to other bins. Games that are not discrete time steps also present a challenge to this methodology. UrbanSim has 15 discrete turns for the player to make decisions. While the player is making decisions the environment is static and does not continue to change. Game scenarios that are continuously create a new timing dynamic for the learner, thus a new dynamic for the scenario designer to consider during design and testing.

C. FUTURE WORK AND RECOMMENDATIONS

This thesis sought to develop a methodology to evaluate ill-defined problem scenarios against their intended training objectives. Through this research other potential research questions were identified.

First, this methodology should be extended to address training objectives that are more specific than strategies or policies and focus on particular actions. The Al Hamra 2 scenario seeks to train students to understand that if one of the two gas stations in the area of operations is damaged, that this should trigger the student to overtly protect the other remaining gas station that is critical to the area of operations.

Second, this methodology should address other fielded UrbanSim scenarios to provide a better understanding of those underlying reward structures. This would provide rather immediate feedback to the user community about the efficacy of the scenarios compared to the intended training objectives.

Third, this methodology should be applied and utilized to develop an entirely new scenario to determine how and when the scenario designer should conduct developmental and formative evaluations. This would serve as an important tool in the overall scenario design process that is not currently available.

Fourth, this methodology should be utilized to assess other scenarios in other games that address ill-defined problems. There are some unique aspects of UrbanSim and PsychSim that may not be present in other games that may provide better insight about the scenario evaluation methodology.

The assessment of training scenarios with respect to the intended training objectives should be formalized for scenario developers at institutional learning centers. Additionally, future simulation and game development efforts should include the capability to assess scenarios with automated tools in the requirements documents to ensure this ability is available and accessible to the training developers.

LIST OF REFERENCES

Clark, R. C. (2008). Building expertise: Cognitive methods for training and performance improvement (3rd ed.). San Francisco, CA: Pfeiffer.

Domeshek, E. A., Holman, E., & Luperfoy, S. (2004). Discussion control in an automated socratic tutor. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, (pp. 1–11). Orlando, FL.

Domeshek, E. (2002). *Technical report 1124 phase i final report on an intelligent tutoring system for teaching battlefield command reasoning skills.* Washington, D.C.: U.S. Army Research Institute.

Domeshek, E. (2004). *technical report 1124 phase ii final report on an intelligent tutoring system for teaching battleifield command reasoning skills.* Washington, D.C.: U.S. Army Research Institute.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in acquisition of expert performance. *Psychological Review*, 363–406.

Fullerton, T. (2008). *Game Design Workshop: A playcentric approach to innovative games.* Oxforc, UK: Morgan Kaufmann.

Intelligent Automation Incorporated. (2011). *Performance Assessment for Complex Simulations.* Orlando, FL: U.S. Army RDECOM-STTC.

Jonassen, D. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional-Design Theories and Models* (Vol. II). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education.* San Francisco, CA: Pfieffer.

Kelley, C., & McLaughlin, A. (2011). Individual differences in the benefits of feedback for learning. *Human Factors: The Journal of the Human Factors and Ergonomics*, 26–34.

McAlinden, P., Durlach, P. J., Lane, H., Gordon, A. S., & Hart, J. (2008). UrbanSim: a game-based instructional package for conducting counterinsurgency operations. *26th Army Science Conference* (pp. 1–11). Orlando, FL: Institute of Creative Technologies, University of Southern California.

McAlinden, R., Gordon, A. S., Lane, H., & Pynadath, D. (2009). *UrbanSim: A game-based simulation for counterinsurgency and stability focused operations.* Marina del Rey, CA: Institute for Creative Technologies, University of Southern California.

Murphy, C. (2011). Why games work and the science of learning. *Interservice, Interagency Training, Simulations, and Education Conference*.

Neill, J. (2012, August 10). *Experiential learning cycles*. Retrieved from http://wilderdom.com/experiential/elc/ExperientialLearningCycle.htm

Ong, J. (2007, November/December). Automated performance assessment and feedback for free-play simulation-based training. *Performance Improvement*, pp. 24–31.

Ong, J., & Ramachandran, S. (2003). *Intelligent tutoring systems: Using AI to Improve Training Performance and ROI.* San Mateo, CA.

Savery, J. R., & Duffy, T. M. (1998). Problem based learning: an instructional model and its constructivist framework. In B. G. Wilson (Ed.), *Constructivist Learning Environments: Case Studies on Instructional Design.* Englewood Cliffs, NJ: Educational Technology Publications.

Smith, R. (2009). *Military simulation and serious games: Where we came from and where we are going.* Orlando, FL: Modelbenders Press.

Stottler Henke Associates. (2012, July 16). *Intelligent tutoring added to BC2010 simulation via HLA interface*. Retrieved from Stottler Henke Associates: http://www.stottlerhenke.com/solutions/training/bc2010_its.htm#

Stottler, R. H., & Vinkavich, M. (2000). Tactical action officer intelligent tutoring system (TAO ITS). *Interservice/Industry Training Simulation Education Conference 2000* (pp. 1–12). Orlando, FL: Interservice/Industry Training Simulation Education Conference.

Stottler, R. H., Jensen, R., Pike, B., & Bingham, R. (2002). Adding an intelligent tutoring system to an existing training simulation. *Interservice Industry Training Simulation Education Conference (I/ITSEC).* Orlando, FL.

Stottler-Henke Associates. (2012, July 16). *ComMentor: Socratic tutoring for high-level command skills*. Retrieved from Stottler Henke Associates: http://www.stottlerhenke.com/solutions/training/commentor.htm

Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning.* Cambridge, MA: MIT Press.

Team Orlando. (n.d.). *Team Orlando*. Retrieved July 7, 2012, from http://www.teamorlando.org/

U.S. Army. (2011). *Army doctrine publication 3-0, unified land ope*rations. Washington, DC: U.S. Army.

U.S. Army. (2006). *Field manual 3-24, counterinsurgency*. Washington, DC: U.S. Army.

U.S. Army. (2011). *Field Manual, FM 7-0, Training units and developing leaders for full spectrum operations*. Washington, D.C.: U.S. Army.

U.S. Army RDECOM. (2011). UrbanSim Training Package. Orlando, FL: U.S. Army RDECOM.

U.S. Army. (2011). *TRADOC PAM 525-8-2 The U.S. Army learning concept for 2015*. Fort Monroe, Virginia: U.S. Army.

USC Institute for Creative Technology. (2012). Retrieved February 16, 2012, from http://ict.usc.edu/projects/urbansim

Wang, N., Pynadath, D., Marsella, S., Cerri, S., Clancey, W., Papadourakis, G. et al. (2012). Toward automatic verification of multiagent systems for training simulations. In *Intelligent Tutoring Systems* (pp. 151–161). Berlin / Heidelburg: Springer.

Wansbury, T. (2011). UrbanSim Project Leader, U.S. Army Research, Development, and Engineering Command. (B. Vogt, Interviewer)

Wansbury, T., Hart, J., Gordon, A. S., & Wilkinson, J. (2010). UrbanSim: training adaptable leaders in the art of battle command. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2010*, (pp. 1–10). Orlando, FL.

Wilson, B. G. (1996). *Constructivist Learning environments: Case studies in instructional design.* Englewood Cliffs, NJ: Educational Technology Publications.

Woolf, B. P. (2008). *Building intelligent interactive tutors: Student-Centered Strategies for Revolutionizing E-Learning.* Burlington, MA: Morgan Kaufmann.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California