Technical Report
1209

Skill Transfer and Virtual Training for IND Response Decision-Making: Project Summary and Next Steps

C.E. Rose
R.M. Seater
A.S. Norige

TBD 2016

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

This material is based on work supported by the Department of Homeland Security Science and Technology Directorate National Urban Technology Security Laboratory under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001.

This report is the result of studies performed at Lincoln Laboratory, a federally funded research and development center operated by Massachusetts Institute of Technology. This material is based on work supported by the Department of Homeland Security Science and Technology Directorate National Urban Technology Security Laboratory under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Department of Homeland Security Science and Technology Directorate National Urban Technology Security Laboratory.

Massachusetts Institute of Technology
Lincoln Laboratory

Skill Transfer and Virtual Training for IND Response Decision-Making: Project Summary and Next Steps

C.E. Rose
A.S. Norige
Group 44

R.M. Seater
Group 45

Technical Report 1209

TBD 2016

Lexington Massachusetts

ABSTRACT

MIT Lincoln Laboratory (MIT LL) has been engaged in a project sponsored by the Department of Homeland Security Science and Technology Directorate (DHS S&T) and the Federal Emergency Management Agency (FEMA) that examines alternative mechanisms for training and evaluation of emergency managers (EMs) to augment and complement existing techniques. The effort seeks to answer the question of how government agencies can ensure that key emergency response personnel have the required skills and knowledge to make critical decisions during an incident of unprecedented size, scope, and complexity, such as an improvised nuclear device (IND) detonation. The effort has examined game-based training methods to determine their suitability for addressing the identified gaps and to improve the tools, techniques, and guidance for creating game-based training materials. This report summarizes the key steps and findings of the project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1. INTRODUCTION

MIT Lincoln Laboratory (MIT LL) has been engaged in a project sponsored by the Department of Homeland Security Science and Technology Directorate (DHS S&T) and the Federal Emergency Management Agency (FEMA) that examines alternative mechanisms for training and evaluation of emergency managers (EMs) to augment and complement existing techniques. The effort seeks to answer the question of how government agencies can ensure that key emergency response personnel have the required skills and knowledge to make critical decisions during an incident of unprecedented size, scope, and complexity, such as an improvised nuclear device (IND) detonation. The effort has examined game-based training methods to determine their suitability for addressing the identified gaps, and to improve the tools, techniques, and guidance for creating game-based training materials. The following three reports resulted from the project research, covering key aspects of the problem space:

- **Analysis of Decision Making Skills for Large-Scale Incidents** [1] summarizes the key decisions and skills that professional EMs thought were critical to an effective response to a large scale incident, with an emphasis on those decisions and skills that may not be adequately addressed by current training methods. The information was gathered through a combination of interviews and surveys conducted with a broad range of Subject Matter Experts (SMEs) and a review of pertinent planning documents. The analysis focused on IND detonation scenarios; however, many of the results generalize to other large-scale incidents.

- **Game Design for Disaster Response Training** [2] provides guidelines for determining when game-based training techniques are suitable to a training task, creating training games for targeting decision-making skills, and validation methods for those games that are cognizant of practical limitations on time and resources. The report also describes two games implemented as part of the project to target skills and decisions identified in the decision-making report mentioned previously. Those two games serve as concrete case studies to demonstrate the guidance on selection, design, and validation.

- **Models for Government-Industry Collaboration for the Development of Game-Based Training Tools** [3] describes the unique challenges of serious game development, how proper use of collaboration partners can overcome those challenges, and how projects can be organized and managed to accommodate those partners. The report presents seven different collaboration models, describes the tradeoffs of each model, and offers guidance on how to select the right model based on the task and budget. The analyses and recommendations in the report are supported by the input of the *Serious Games Focus Group* held at the Massachusetts Institute of Technology on August 27, 2015.

The findings of those reports are summarized in Section 2; the full reports are available upon request. Section 3 focuses on the practical next steps to apply the findings of this project.

2. PROJECT SUMMARY AND TAKEAWAYS

2.1 EMERGENCY MANAGEMENT TRAINING FACES CHALLENGES

Adept EMs and incident commanders possess a refined set of behavioral competencies, such as the ability to improvise, make correct inferences under time pressure, and provide effective guidance to teams. It is difficult and time-consuming to gain expertise in those vital decision-making roles, especially for large-scale incidents. The required skills are challenging to learn with traditional training methods, such as classroom lectures or live exercises, and opportunities to gain actual experience are infrequent. Classroom-based training provides professionals with the fundamentals of their position requirements, but it has a limited ability to train or assess strategic thinking and interpersonal skills. Live exercises help to rehearse knowledge, but they are time-consuming and generally expose the participant to just a single scenario. On-the-job learning is invaluable for building and demonstrating decision-making skills, but it is fundamentally limited by the infrequency of relevant events in the real world. Waiting for candidate emergency managers to slowly acquire and demonstrate the skills during real incidents limits the available talent pool and reduces confidence that decision makers have the skills to handle large incidents.

In this project, we began with identifying the difficult decisions that EMs would be required to make during an IND detonation or similar large-scale incident. We then investigated the skills that experts used when making those impactful decisions with an emphasis on which skills were most difficult to train or evaluate. These difficult decisions and required skills were identified through a quantitative and qualitative domain analysis of relevant technical documents and input from a broad range of SMEs. While we focused our efforts in one area of disaster response, a similar analysis could be performed to identify training difficulties for any domain. The top ranked decisions and skills relevant to an IND detonation or other large-scale incident are illustrated in Figure 1.

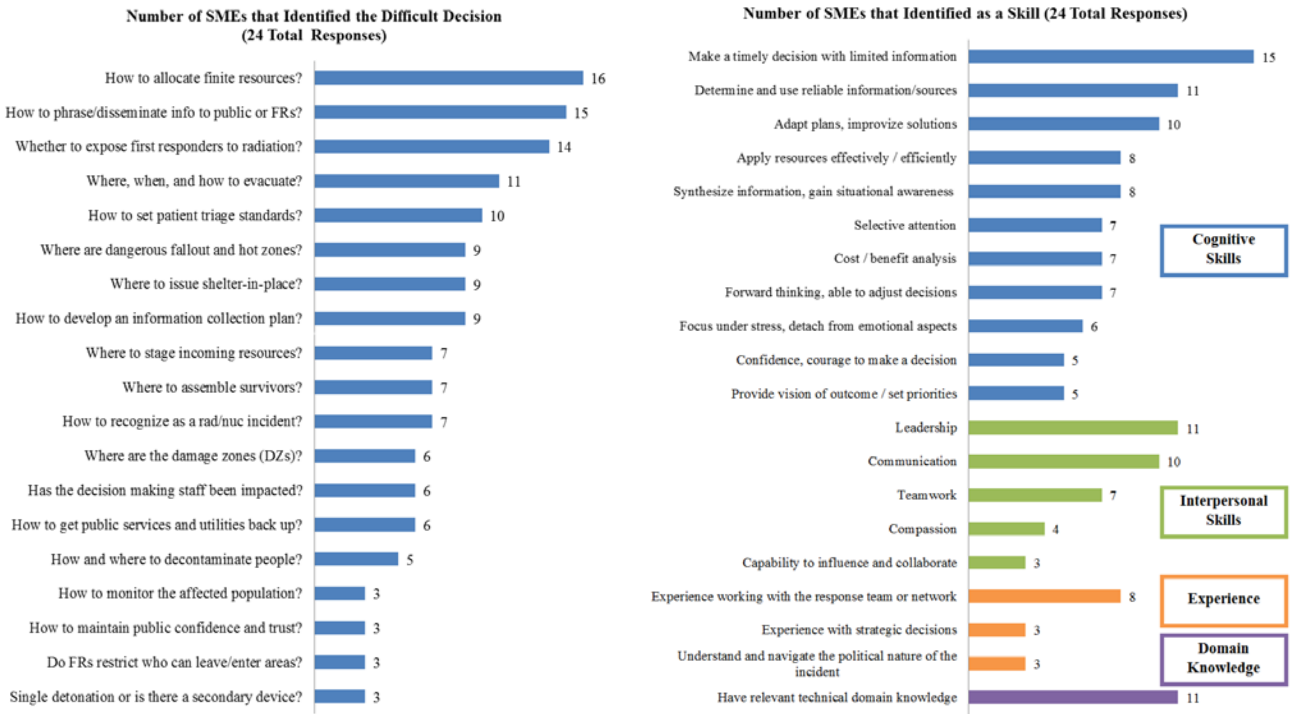

Figure 1. Difficult Decisions (left) and Critical Skills (right) ranked by their frequency of emphasis by SMEs.

2.2 GAME-BASED TRAINING CAN PLAY AN IMPORTANT ROLE

Game-based training for emergency management professionals can complement traditional training methods such as classroom learning, live exercises, and on-the-job experience. It offers a number of appealing benefits to traditional methods, such as reduced cost, greater accessibility, flexible scenario creation, and increased motivation for participation. It can also effectively engage participants in a variety of complex decision-making environments that are not easily replicated by other training methods.

The use of serious games has attracted a growing interest in both the research and training communities, and the related field of "war gaming" has a long tradition in military sectors. Games and game-like training methods are already used throughout the Department of Homeland Security (DHS) and other government agencies to support training, evaluation, analysis, and technology exploration. Those techniques have found successful niches, but their wider applicability faces several barriers that limit their impact. This project offers practical guidance to government agencies, project offices, and program managers on how to overcome those obstacles and leverage the strengths of game-based training methods.

Two prototype games developed as part of this program are illustrated in Figure 2, showing two very different styles of games that can target decisions and skills key to effective emergency management. *First Response* demonstrates how a strategy game can target topics related to allocation of scarce resources, managing competing priorities, and adjusting the response strategy as the situation unfolds.

Disaster Dilemma demonstrates how a narrative game can target topics related to managing social and political situations that emerge in parallel to the physical disaster, and how to make calculated tradeoffs based on shifting context, priorities, and opportunities.

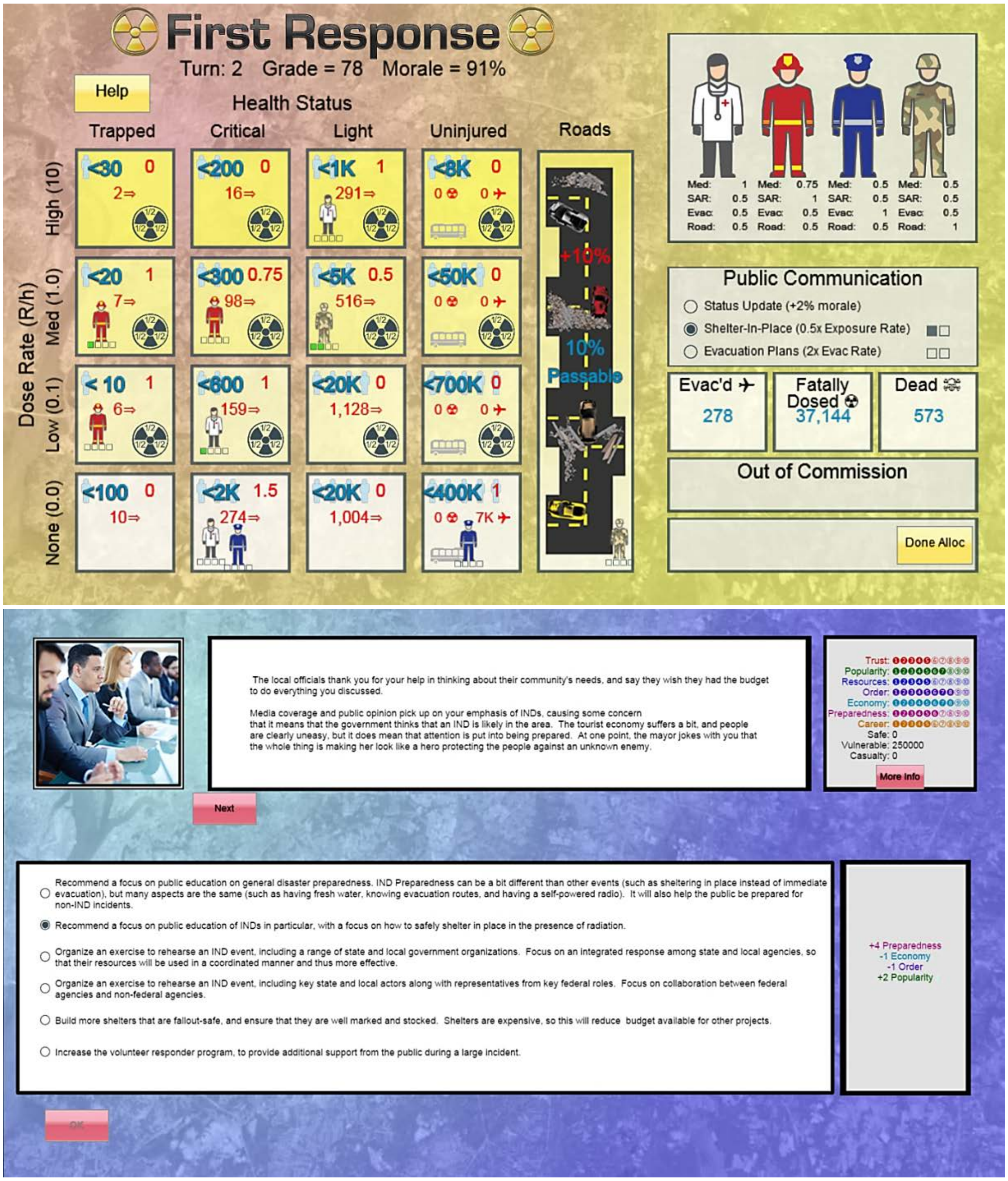

Figure 2. Two lightweight serious games (LWSGs) developed to target skills important to emergency management. First Response (top) and Disaster Dilemma (bottom).

2.3 WHEN IS GAME-BASED TRAINING THE RIGHT APPROACH?

Serious games are interactive systems that focus the participant on decision making in a virtual scenario. They offer players rapid feedback, multiple play iterations, a low cost of failure, and an opportunity to experiment and explore different strategies. They are well-suited for building and evaluating strategic and interpersonal skills, and can be used in concert with classroom lectures (which are effective at conveying factual information) and live exercises (which are effective at rehearsing complex procedures).

Different types of serious games can serve different audiences and different types of training materials, as illustrated in Figure 3. Quiz Games are best for conveying factual information, Tactical Games help convey difficult tradeoffs and depend on the decision context, and Cognitive Games capture subtleties about situational awareness, information processing, and prioritization.

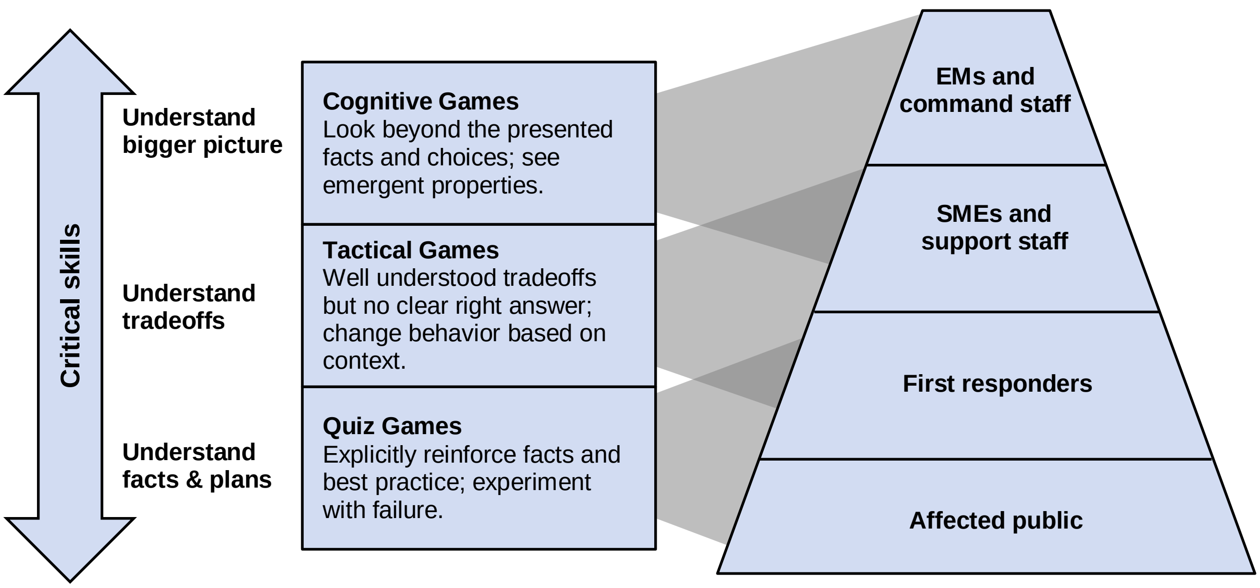

Figure 3. Different types of learning objectives, different styles of serious games, and different audiences are related.

2.4 ADVANTAGES OF LIGHTWEIGHT GAMES

Conventional approaches to using game-like structures for serious purposes often encounter several limitations due to being heavyweight. Heavyweight approaches have high up-front development costs, high scenario creation costs, a high burden on attendees' time, they give players limited exposure to a variety of scenarios, and they rarely yield quantitative results. In contrast, lightweight serious games (LWSGs) focus on having **short** playtime (minutes, not days), **deployable** implementations (playable over the Internet, not requiring specialized hardware), and **focused** in content (a few learning objectives, not a full simulation). As illustrated in Figure 4, players can experience many more scenarios and iterations of those scenarios, providing more reinforcement of learning objectives, more opportunity to experiment with different strategies, and more opportunity to gather statistically significant data on their readiness and gaps.

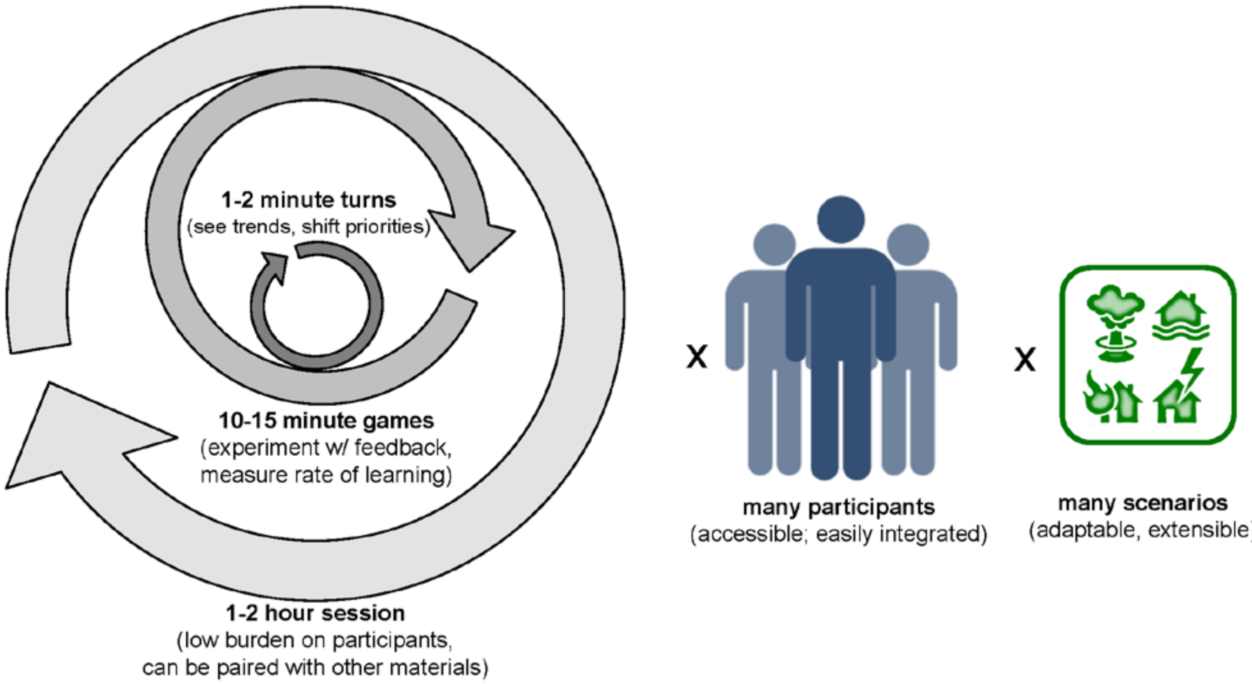

Figure 4. LWSGs are short, accessible, and focused.

2.5 GAME-BASED TRAINING VALIDATION

Serious-game validation is the most important step during the design process of any training method, game-based or otherwise. It is in this step that we justify that the created game targets the appropriate technical material, lessons, decisions, and skills that were intended by the game creators. The most straightforward approach to validation is to run a direct end-to-end experiment and compare on-the-job performance of people who do and do not play the game. However, traditional empirical user studies can often be infeasible, impractical, or otherwise undesirable for a variety of reasons. They are generally costly, require a long time frame, and can produce noisy results because of confounding factors. Scientific field studies are essential for establishing the validity of game-based training in general, but they are not well-suited for supporting the creation of games for particular applications. This project described and demonstrated the utility of several alternate methods for building confidence in the suitability of a particular game for a specific training task. The methods explored by this project include using frequent SME involvement, following best practices in game design, analyzing expert strategies, and using machine learning and artificial intelligence (AI) techniques.

2.6 FORGE NEW PARTNERSHIPS TO SUPPORT GAME-BASED TRAINING DEVELOPMENT

Creating game-based training tools for emergency management presents unique challenges as game design often requires multidisciplinary contributors to develop a successful product. The diverse set of stakeholders, which includes the sponsoring government agencies, game developers, and those who may benefit from the training, necessitates collaboration between multiple organizations, which can lead to a challenging management and development process. This program investigated the practical question of how to leverage existing talent in the commercial, academic, and federally funded research and development center (FFRDC) communities to help interested government agencies create effective game-based training tools. The analyses and recommendations in this area are supported by the input of the *Serious Games Focus Group* held at the Massachusetts Institute of Technology in August 2015.

Effectively collaborating with new partners may require adapting existing collaboration models to align with the interests and working patterns of those communities and supporting the inherently iterative process of creating an effective training game. We recommended seven different collaboration models, all of which have precedents both in and outside of government programs. An appropriate model can be selected based on the needs and budget of the program; a summary is illustrated in Figure 5.

	Model	Approx. Project Budget	Initial Setup	Availability of Opportunities	Likelihood of Impact	Product Maturity	Similar Government Structures	Commercial Precedent
A	**Traditional Project Sponsorship**	$500k-10M	None	On Demand	Low - Med	High	Sole Source Project	Mission US, Mindset Quest, Triad Interactive
B	**Game Design Competition**	$100k-200k	None	On Demand	Med - High	Med	DARPA Challenge, BAA	Global Game Jam, Game Jam Central, IndieGameJams.com
C	**Project Bidding and Development Platform**	$10k-100k	Required	On Demand	Low - Med	Med	BAA	ArtCorgi, Bricklink, Upwork, Toptal, Shapeways
D	**Academic Capstone**	$10k-100k	Required	On Demand	Low - Med	Low	Existing Centers of Excellence	MIT Beaverworks, NASA Glenn Research Center
E	**Commerical Dual Release**	$1M-10M	None	Require Opportunity	Med	High	Sole Source Project	Full Spectrum Warrior, America's Army, NumberShire, Sokikom
F	**Special Game Mode**	$100k-1M	None	Require Opportunity	Med	High	Sole Source Project	League of Legends, Plants vs. Zombies 2, Quake, Portal 2
G	**Center of Excellence**	$100k-10M (varies by sub-model)	Required	On Demand	High	Med	Existing Centers of Excellence, FFRDCs	MassDigi, CIGG

Figure 5. Comparison of the strengths and weaknesses of the seven collaboration models.

3. NEXT STEPS

There are several productive ways to build on the findings of this project, depending on the goals and interests of the sponsoring agencies. These extensions fall into three main categories: applied pilot studies, collaboration building, and validation improvement.

3.1 APPLIED PILOT STUDIES

We recommend performing initial integrations of game-based training with courses at current training programs for emergency professionals, such as those run by FEMA for Incident Management Assistance Team (IMAT) training, federal training organizations such as the FEMA Emergency Management Institute (EMI), or local emergency management offices. Preliminary conversations suggest that many such organizations are open to integrating serious games to complement their current methods, and several have expressed eagerness to experiment with game-based training. Running integrated studies with working training organizations is likely to yield more actionable feedback to improve the training methods and build more confidence in their broader merit than would additional theoretical or laboratory studies.

We recommend starting with an integrated approach, where a classroom or online training course is paired with a game targeting similar learning objectives. The students would play the game before the course to familiarize themselves with the challenges and tradeoffs they face. Then, the students would participate in the normal course, with some context for the lessons they are being taught. After the course, the students would play the game again to solidify the material they just learned, experiment with variations of the material, and receive some repetition of the material to improve retention. If the course is multi-day, the students might play the game each day after the lecture and discuss it with other students.

As with any experiment, to understand the benefit of game-based training it is important to have a controlled comparison. We suggest creating three groups of students to participate in analyzing the benefits of a course with game-based training integration. One test group would not play the game and take the training course in its current state; one test group would take the course with the game integrated; a final test group would take the course, spending the same amount of time as the game-playing group, reviewing a summary of the lecture material. In this way, researchers can compare performance of the game versus nothing and versus additional time with the material. Note that even if the benefit gained from the game-based training is the same as spending additional time with the material, that is a positive result for serious games—one of the benefits of serious games is that they engage players, resulting in more total time spent with the material (in addition to showing them that material in a more interactive format). Running such a study to would offer several benefits and opportunities:

- **Quantitative Validation**. Have students of all groups take a quiz to evaluate their comprehension of the material. Compare the results of the quiz for the groups of students taught via traditional methods versus when paired with a serious game.

- **Qualitative Validation**. Ask students and instructors to self-report on what benefits, drawbacks, and opportunities for improvements they observed. Such feedback would help to improve the games, adjust the way in which the game is incorporated into the curriculum, and offer some basic validation of the merits of the game.

- **Professional Validation**. Run the experiment for a live classroom lecture with a discussion component. Do not tell the instructor which students were using the game-based methods and which students were just seeing the lecture. After the course, have the instructors rate their assessment of the depth of understanding that each student had by the end of the course.

- **Retention Study**. Several weeks or months after the instructional session, have the students re-take a quiz or provide another self-assessment of their knowledge of the material. Compare the retention of material when taught just via traditional methods versus when paired with a game.

- **Compare Impact of Experience**. In addition to having students play the game, give the game to experienced professionals. Determine what professional experience and demographic factors affect performance on the game to identify what types of material the game is most effective at capturing. For example, if experienced IMAT outscore health physicists, that provides evidence that the game is more directly targeting decision-making skills and not factual knowledge about radiation. If experienced responders outscore inexperienced volunteers, then there is evidence that the game is measuring skills that are acquired through experience.

- **Measure Topics Discussed**. Have students play the game cooperatively in small teams of 2–4 players. Record the discussions they have about what actions to take in the game, and code them according to topics (e.g., “inferring priorities,” “debating rules,” “mathematical calculations,” “planning for evacuation”). Such a measure would determine if the game is forcing players to think about the topics it was intended to address.

- **Demonstrate Template Flexibility**. Integrating one of the games developed during the current project [2] will involve updating the game to meet the particular topics and learning objectives of the collaborating training organization. Doing so will help improve understanding of when different game styles and templates mesh well with different topics, producing better guidelines for when different types of games are most appropriate.

- **Explore Integration Styles**. Compare different ways of pairing game-based training with traditional lectures and coursework. For example, how do students rate the game differently if they play both before and after the lecture versus just one of those times?

- **Student Benefit**. There is also an immediate benefit to the students involved, who get early exposure to modernized training methods and see the course material from another angle.

3.2 BUILD COLLABORATIONS

Partnering with new collaborators in order to build, validate, and deploy serious games may require modifying existing models for collaboration. Two communities that are likely to be very productive partners—independent video-game developers and academic game degree programs—are not familiar with working with government sponsors. Building those connections through small efforts can pave the way and produce a smooth relationship so that those partnerships are available for later efforts.

- **Academic Capstone Projects**. During the last decade, university degree programs have sprung up that are dedicated to training students to enter the game industry. Such programs typically combine computer science and art coursework, and they often have a very practical focus on providing students with skills to immediately start a game-development career, and the professors in those programs typically combine practical experience with deep theoretical thinking. Engaging such programs through capstone projects or as centers of excellence can be an inexpensive way to access a smart community that is able to manage the link between best practices in game design and scientific principles of user analysis of experimental validation.

- **Independent Video-Game Developers**. Small companies and individuals that pursue video-game design and development can be inexpensive sources of talent, creativity, and productivity. However, that community is unfamiliar with operating with government agencies via traditional methods and is not equipped with the administrative infrastructure to engage government bureaucracies. As recommended in one of our reports [3], one way to engage those groups is through game design competitions and "game jams," akin to a Defense Advanced Research Projects Agency (DARPA) Challenge. Running a competition to create a serious game for a specific emergency management training program would be a good way to evaluate the strengths and weaknesses of games coming out of that collaboration mode, and would likely produce either usable games or strong prototypes to directly address a training need.

3.3 IMPROVE VALIDATION TECHNIQUES

As discussed in our report [2], a serious game can be partially validated by analyzing winning strategies. If winning strategies involve the types of decisions and choices that we would like to instill in students, then the game is rewarding desirable behavior and thus more likely to be teaching that behavior. In contrast, if a winning strategy involves undesirable or unrealistic behavior on the part of the player, then we know that the game needs to be revised to avoid teaching the wrong lessons. There are two extensions to this work that would help to make that style of validation easier and more trustworthy.

- **AI-Based Validation**. Work should be done to improve machine learning and AI-based methods for determining optimal strategies in a game. These techniques are often applied to games for the purpose of creating computer opponents for entertainment games, but additional research is needed to improve their use as a tool to support the design, refinement, and validation of serious games. Once available, these techniques could become low-cost ways to build confidence that a game is teaching the right lessons, and not accidentally rewarding undesirable behavior.

- **Compare Transference and Precision**. A transference test measures if students using a method (e.g., serious games) are more effective in their real-world jobs, but does not determine if the right lessons are being taught. A precision test measures if the right lessons are being taught, but does not measure if that translates to better job performance. It would be interesting to do a comparison study to determine how well precision tests predict transference. For example, choose a game that rates well according to the alignment of winning strategies with learning objectives (via the methods described in our report [2]), and run end-to-end tests of transference to see how well precision predicts transference. If precision does predict transference, then less expensive validation methods can be employed to build and validate game-based training materials with confidence that those materials will have a positive operational impact, even though such impacts are hard to directly measure.

GLOSSARY OF ACRONYMS

AI	artificial intelligence
CBRNE	chemical, biological, radiological, nuclear, and high-yield explosive
DARPA	Defense Advanced Research Projects Agency
DHS	Department of Homeland Security
DHS S&T	Department of Homeland Security Science and Technology Directorate
EM	emergency manager
EMI	Emergency Management Institute
FEMA	Federal Emergency Management Agency
FFRDC	federally funded research and development center
IMAT	Incident Management Assistance Team
IND	improvised nuclear device
LWSG	lightweight serious game
MIT LL	MIT Lincoln Laboratory
OEM	Office of Emergency Management
SME	subject matter expert

WORKS CITED

[1] C. Rose, R. Seater, and A. Norige, "Skill Transfer and Virtual Training for IND Response Decision Making: Analysis of Decision Making Skills for Large-Scale Incidents," MIT Lincoln Laboratory, Lexington, MA, Technical Report 1206 (insert publication date when final).

[2] R. Seater, C. Rose, A. Norige, J. McCarthy, M. Kozar, and P. DiPastina, "Skill Transfer and Virtual Training for IND Response Decision Making: Game Design for Disaster Response Training," MIT Lincoln Laboratory, Lexington, MA, Technical Report 1207 (insert publication date when final).

[3] R. Seater, C. Rose, A. Norige, K. Chang, P. DiPastina, and J. McCarthy, "Skill Transfer and Virtual Training for IND Response Decision Making: Models for Government-Industry Collaboration for the Development of Game-Based Training Tools," MIT Lincoln Laboratory, Lexington, MA, Technical Report 1208 (insert publication date when final).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* TBD 2016	**2. REPORT TYPE** Technical Report	**3. DATES COVERED** *(From - To)*

4. TITLE AND SUBTITLE

Skill Transfer and Virtual Training for IND Response Decision Making: Project Summary and Next Steps

5a. CONTRACT NUMBER
FA8721-05-C-0002 & FA8702-15-D-0001

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Charles Rose, Robert Seater, and Adam Norige

5d. PROJECT NUMBER
9-044401

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

TR-1209

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Homeland Security
National Urban Security Technology Laboratory
201 Varick Street
New York, NY 10014-9444

10. SPONSOR/MONITOR'S ACRONYM(S)
DHS S&T/NUSTL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

MIT Lincoln Laboratory (MIT LL) has been engaged in a project sponsored by the Department of Homeland Security Science and Technology Directorate (DHS S&T) and the Federal Emergency Management Agency (FEMA) that examines alternative mechanisms for training and evaluation of emergency managers (EMs) to augment and complement existing techniques. The effort seeks to answer the question of how government agencies can ensure that key emergency response personnel have the required skills and knowledge to make critical decisions during an incident of unprecedented size, scope, and complexity, such as an improvised nuclear device (IND) detonation. The effort has examined game-based training methods to determine their suitability for addressing the identified gaps and to improve the tools, techniques, and guidance for creating game-based training materials. This report summarizes the key steps and findings of the project.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Same as report	20	**19b. TELEPHONE NUMBER** *(include area code)*

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