Using Games to Accelerate Aircrew Cognitive Training

Dr. Tricia Mautone
Anacapa Sciences, Inc.
Santa Barbara, CA
tmautone@anacapasciences.com

Dr. V. Alan Spiker
Anacapa Sciences, Inc.
Santa Barbara, CA
vaspiker@anacapasciences.com

Dr. M. Ron Karp
Arizona State University
Phoenix, AZ
ron.karp@asu.edu

Dr. Curtis Conkey
Naval Air Warfare Center Training Systems Division
Orlando, FL
curtis.conkey@us.army.mil

ABSTRACT

Game-based training (GBT) is a promising medium for increasing the efficiency of training complex cognitive skills and improving performance. However, there is little empirical research to guide decisions about when, or whether, GBT should be integrated into a professional training curriculum. To fill this void, we developed a rule-based decision tool that helps identify which training areas would likely benefit from the insertion of GBT technology, and then makes recommendations about which game elements and design patterns (e.g., type of challenges, feedback format, etc.) would be most appropriate for games targeting those skill areas.

To develop and test the tool, we conducted an analysis of the undergraduate pilot training program at Arizona State University, identifying flight management system (FMS) operation as a good candidate for GBT. Using our analysis of the task environment, we then selected potential best-fit game elements to develop a fully-functional web-based game to train student pilots how to program an aircraft's FMS quickly and accurately. The game's effectiveness was then evaluated in a series of studies in which half the students received the FMS game and half received conventional computer-based training. All students then took transfer criterion tests, using a simulated FMS device. On near and delayed transfer tests, students who received GBT scored significantly higher (making fewer errors and omissions) than their counterparts who received conventional training ($p<.05$).

Using lessons learned from our FMS game development and evaluation, we then applied the GBT tool to create and evaluate a new training game for a different domain, aircrew communication. The paper provides an analysis of the FMS and crew communication game designs, the quantitative results of the criterion evaluations, and provides a roadmap for how to facilitate the development of effective training games by migrating proven GBT design patterns to comparable training applications.

ABOUT THE AUTHORS

Tricia Mautone, senior scientist at Anacapa Sciences since 2003, primarily focuses on developing and evaluating effective game-based and multimedia instruction. She earned a Ph.D. in cognitive psychology from the University of California, Santa Barbara and has several publications on multimedia instruction. She has presented at numerous conferences, including ITEC, AERA, SERP, and IITSEC.

V. Alan Spiker has been a principal scientist at Anacapa Sciences since 1982. He is responsible for human factors research associated with advanced technology and has conducted field studies on aviation training and performance for industry and all branches of service. Dr. Spiker received a Ph.D. in Experimental Psychology from the University of New Mexico in 1978.

Merrill R. (Ron) Karp is Professor Emeritus with the Department of Aeronautical Management Technology, Arizona State University, Polytechnic Campus, where he developed and managed the ASU/Mesa Airlines Airline Bridge Training Program. Dr. Karp received his Ph.D. in Administration & Management from Walden University in 1996. He served as a Fighter Wing Commander of F-4G Wild Weasels in Operation Desert Storm and flew the F-16 as a Wing Vice Commander and Director of Operations. He also served three combat tours in Vietnam.

Dr. Curtis Conkey from Naval Air Warfare Center Training Systems Division (NAWCTSD) is the Principle Investigator and Team Lead for the Learning Technologies Lab. Mr. Conkey has prior experience in conventional simulation systems having worked with Onboard Team Training applications for the Next Generation Virginia class submarine. Dr. Conkey received his Doctorate in Modeling and Simulation from the University of Central Florida.

Using Games to Accelerate Aircrew Cognitive Training

Dr. Tricia Mautone
Anacapa Sciences, Inc.
Santa Barbara, CA
tmautone@anacapasciences.com

Dr. V Alan Spiker
Anacapa Sciences, Inc.
Santa Barbara, CA
vaspikcer@anacapasciences.com

Dr. M Ron Karp
Arizona State University
Phoenix, AZ
ron.karp@asu.edu

Dr. Curtis Conkey
Naval Air Warfare Center Training Systems Division
Orlando, FL
curtis.conkey@navy.mil

INTRODUCTION

A combination of factors is placing increased pressure on military and industrial training communities to look for innovative ways to maximize training efficiency and effectiveness while reducing training costs. With challenges ranging from reductions in manning, to implementation of new technology systems, to changing minimum-experience requirements for incoming trainees, there is an expanding emphasis on finding ways to select the optimal mix of training methods that best utilize training resources, increase trainee throughput, and reduce trainee attrition.

This is particularly true with aircrew training, where the requisite knowledge/skills/attitudes (KSAs) are complex and varied, the stakes for success and failure are high, and the training itself – much of it simulator-based – is quite expensive. In this environment, training gaps, whether due to lack of prerequisite knowledge, insufficient practice, or skill decay, often result in less-than-desirable learning outcomes, valuable time and resources spent on remedial training, or in high washout rates, all of which contribute to increased training costs (Salas, Wilson, Priest, & Guthrie, 2006).

A promising method for helping to bridge many of these training gaps is the integration of Game-Based Training (GBT) at key locations within an organization's training curriculum or Program of Instruction (POI). If used appropriately, GBT has the potential to heighten learner motivation; provide the opportunity for applied, goal-oriented, repetitive practice; and allow for feedback-rich, structured self-study – all of which should ultimately lead to improved training outcomes and greater training efficiency (Mautone, Spiker, & Karp, 2008). Despite this potential, there are presently few empirical studies that provide clear-cut data on the effect of GBT on training outcomes and how this method compares to more conventional training methods. Nor are there many empirically-supported guidelines upon which to make informed decisions about where or when the application of GBT is likely to be effective – or on how to design or select game features that best meet the needs and constraints of a particular training environment.

In this paper, we describe the research from a NAVAIR-funded project whose primary objectives are to: (1) gather well-controlled empirical data comparing the effectiveness of GBT to more conventional training methods (e.g., non-game computer-based training or static simulator practice) within an actual "real-world" training environment – in this case Arizona State University's Professional Flight Training Program; and (2) develop and validate a systematic, science-based approach to selecting and developing games and specific game elements that are best suited to meet the needs of a particular training environment. The approach, and the data gathered during the validation process, will form the foundation for the development of a theoretically-driven and empirically-supported decision-aid tool that specifies when and how specific gaming elements should be integrated into training and instruction. The tool would allow instructional designers, and others, in a variety of domains to make informed decisions about the best uses of GBT to enhance learning and performance outcomes.

THE CASE FOR GAME-BASED TRAINING

Consider the following three scenarios:

- An airline pilot trainee is scheduled for a 30-minute session in a $200/hour high-fidelity simulator. With its sophisticated instruments, and displays, the device is intended to provide pilots with practice flying a commercial aircraft in both routine and emergency situations. However, during this particular session, the trainee never gets to actually "fly the sim;" instead a good chunk her valuable simulator time is spent attempting to input required preflight data into the aircraft's Flight Management System (FMS) while the instructor provides remedial instruction. The problem is not that the trainee is unmotivated or is incapable of performing the task; rather, considerable time has elapsed since she last used the FMS, and she did not have an opportunity to receive the necessary varied and structured hands-on practice with this hard-to-use device.

- A Navy LCAC (Landing Craft, Air Cushion) navigator trainee washes out after several weeks of initial qualification training. He did fine when learning the individual tasks, but did not put in the time or effort needed to practice these skills to the point where he can perform them quickly and effortlessly. As a result, when he begins crew coordination training and has to apply those skills in the simulator, he is unable to handle the heavy multitasking demands and drops out of the program.

- In a course designed to provide Marines with enhanced observation and behavior profiling skills, the instructor requests 100 hours for the training course, but is only allocated 40. As a result, a number of key profiling KSA cannot be developed to the highest levels desired during the course; the instructor is looking for ways to provide bridge training on select skills once the course is over. Pre-training or "sustainment training" in other venues subsequent to the course could also help make better use of limited class time.

As these scenarios illustrate, training on technical and tactical tasks is often hampered by a lack of opportunity to engage in structured practice and/or a reluctance to practice intrinsically unmotivating tasks, which translates into less time-on-task, more time needed to reach criterion levels of performance, inefficient use of resources, and an increased risk of skill decay. Likewise, training in non-technical areas such as decision-making, situational awareness, team coordination, and cultural awareness, to name a few, is often difficult, time-consuming, and expensive to coordinate and effectively execute.

In order to address these challenges, innovative methods are needed that sustain student interest, promote repetitive practice, accelerate learning, provide opportunities for self-study and skill refresher training, and capitalize on trainee downtime. This is where GBT comes in.

The term "Game-based Training" has varied uses; we define GBT as the insertion of games – typically called "Serious Games" or "Immersive Learning Simulations" (Conkey, 2009) – into a POI in order to meet a specific training need. As with the two games used in our studies, the games are often computer-based, though that is not a requirement. What defines a serious game is that, like "regular" games, it (ideally) provides players with a challenging goal, is fun to play, and incorporates rules and scoring. But unlike a regular game, a serious game has been designed specifically for training and, as such, has defined learning outcome objectives, real world relevance (i.e., imparts knowledge and skills that can be applied in the real world), and incorporates integrated instructional support (Aldrich, 2005; Bergeron, 2006; Conkey, 2009; Garris, Ahlers & Driskell, 2002; Mayer, 2003; Prensky, 2001).

The premise behind GBT is that games provide learners with the opportunity and motivation to actively practice and apply important skills in a feedback-rich, situated learning environment, which then results in more focused time on task, better learning outcomes, better transfer to the real world and, ultimately, a more efficient and effective use of training time.

However promising GBT may be, it is it not necessarily a panacea – and one size does not fit all. When considering the integration of GBT into a given training environment, two key questions need to be addressed:

1) **Is GBT the best fit for this training environment?** – i.e., is it an appropriate method for training the targeted KSA? Is it likely to offer any advantages over more conventional training, for *these* learners on *this* task in *this* training situation?
2) **If so, how should GBT be implemented?** – i.e., what *type* of game would be most effective for this training situation? How should the game be structured? What is the optimal selection of feedback formats, challenge types, levels of autonomy and exploration, degree of fantasy and role-play, etc?

To address these questions, we developed a systematic, science-based approach to guide the design and evaluation of effective training games tailored to meet the needs of a specific training environment. This approach serves as the roadmap for the development of our rule-based, web-enabled decision-aid tool, called TARGET (Tool to Apply Robust Gaming Elements to Training), which is designed to provide instructional designers and others with empirically-supported guidelines on when and how to best implement GBT.

In the next section, we provide a brief overview of the GBT design methodology that underlies TARGET. We then describe how we applied this approach to an actual training environment – ASU's Professional Flight Training Program – to first select areas we believed would benefit most from the insertion of GBT, and then to create web-based games that train pilots how to operate an aircraft's flight management system (FMS) and to execute the crew coordination required for flight profiles and callouts. We next describe how the games were integrated into ASU's training program and summarize the results of a series of empirical studies that compared the effectiveness of GBT to more conventional training. We conclude with how the lessons learned from these studies were then used to guide further development of TARGET.

OVERVIEW OF THE TARGET MODEL

Conceived as a web-based decision aid tool that facilitates the selection and development of effective GBT materials, TARGET is based on a systematic approach to serious game design where the results of a guided analysis of relevant characteristics of a candidate training environment are linked via a relational database to best-fit game element variations and design patterns to create a game that can be embedded into an existing training curricula to enhance training outcomes.

The main premise behind the TARGET model is that the benefits of GBT can be maximized – and the costs and risks associated with game development can be reduced – if decisions about when and how to implement GBT are based on empirical evidence and guided by a systematic approach to serious game design. Figure 1 provides a schematic illustration of the main elements and processes of the model.

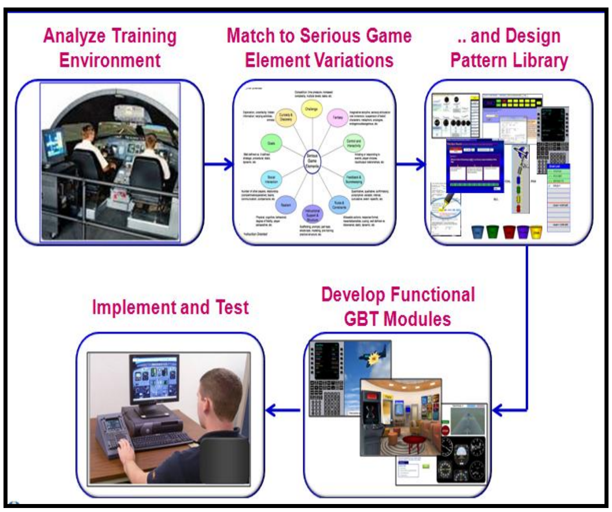

Figure 1: The TARGET Model

The first component of the TARGET model (depicted in the top left-hand corner of Figure 1), is a ***guided analysis of the training environment***. The analysis focuses on distilling parameters of the training environment most relevant for (1) determining whether the candidate training area is likely to benefit from insertion of GBT, and (2) making decisions about how to design the best-fit game. In the web-based version of TARGET, this analysis will be streamlined by using a structured query framework where users of the tool (e.g., instructional designers) are asked a series of branching questions about the targeted training area. Table 1 lists some of the main categories of information gathered in the analysis.

Table 1: Categories of Information Collected in the Guided Training Environment Analysis

Targeted training objectives and associated KSAs - *e.g., Equipment operation, resource management, repair and troubleshooting, crew coordination and communication, monitoring imagery-based data, etc.*
General task characteristics - *e.g., degree of task structure; speed/accuracy tradeoff; number of response options, cue predictability, level of autonomy, etc.*
Underlying cognitive functions involved - *e.g., visual attention, critical thinking, LTM retrieval, classification, planning, etc.*
Type of learning required - *e.g., procedural, conceptual, problem solving, systems, factual, etc.*
Learner characteristics - *e.g., level of prior knowledge, motivation, learning styles, etc.*
Problems with current training (general and specific) - *e.g., lack of opportunity to practice, lack of motivation, gaps in component knowledge, etc.*
Learning environment - *e.g., available training time, anticipated level of instructor involvement, available material resources, location of training, etc.*

Information gathered in this analysis is first used to create a training environment profile, which is then ***matched with data stored in TARGET's relational database***. The database is populated with findings from serious games research, as well as recommendations based on instructional design best practices and on cognitive learning theory – and is organized around the Serious Game Element Taxonomy we developed in earlier phases of our research (see Mautone, Spiker, & Karp, 2008). Table 2 lists five of the ten categories from the taxonomy and provides examples of possible game element variations for each.

Table 2: Selected Serious Game Elements

Challenge - e.g., competition against timer, self, and/or others; number of simultaneous demands; introduction of unexpected events, etc.
Feedback & Scoring - e.g., prescriptive or descriptive; quantitative or qualitative; immediate or delayed, etc.
Rules and Constraints – reward and penalty structure; rules given upfront or discovered; etc.
Structure and Instructional Support – whole task/part-task; expert modeling, faded scaffolding, etc.
Fantasy – depth of storylines, degree of role-playing, sensory stimulation, suspension of belief, etc.
Control and Interactivity – degree of player's control of event outcomes, pacing, criteria for increased autonomy, etc.

Based on criteria stored in the database, TARGET first makes a recommendation about whether GBT is appropriate for the particular training environment. The model might suggest, for example, that GBT might *not* be the most appropriate method for instruction in this situation if learner motivation is not an issue, or the to-be-learned material is fairly straightforward and could be more readily imparted via text or lecture. But if the training area *is* identified as a suitable candidate for GBT, the relational database then ***matches characteristics of the training environment with best-fit game element variations*** (e.g., method of feedback, level of fantasy, type of challenges, etc.) that have been shown to be well-suited for that environment. Matches will, again, be based in part on empirical findings and best-practices instructional design.

The relational database is also linked to the tool's library of reusable ***design patterns***, a concrete example that can be incorporated into another game, along with associated implementation specifications (Mautone, Spiker & Dick, 2007). The final output is a set of recommended game elements and design patterns, which can be used to develop the game.

Once the game is developed, it is then ***implemented*** in the actual training environment and data is gathered regarding the effectiveness of the game and game elements. This data can then be fed back into the relational database to provide additional ***validation of the recommendations***, and to further expand and refine the database.

To illustrate this process, we provide concrete examples of how this approach was applied to an actual training environment.

APPLICATION TO AIRCREW TRAINING Part 1: FMS PROGRAMMING

In order to test and validate our TARGET model, and to gather data for our relational database, we applied this systematic approach to address training needs within ASU's Professional Flight Training Program.

ASU's Aeronautical Management Technology (AMT) department provides its undergraduates with an intensive, four-year program of focused study in aviation that prepares its graduates to fly with US regional and major airlines. Our analysis primarily focused on the training that students receive in their fourth year of study. At this point in their training, students have completed nearly all of their general aviation courses and are now transitioning to commercial aircraft flight training, much of which requires integration of many component skills and interaction with advanced medium- and high-fidelity simulators. Our goal in this analysis was to select areas of the training that our model predicted would most likely benefit from GBT. The first competency we selected was training student pilots how to program the flight management system (FMS) on a Canadair Regional Jet (CRJ) aircraft. Below, we briefly describe FMS operations and the criteria our model used to identify this area as a good candidate for GBT. We then describe how we developed the FMS Game using the TARGET methodology.

Analysis of the FMS Training Environment

The FMS is the pilot's primary interface to the software that controls the plane's navigation and performance. The system is responsible for flight planning, control of navigation sensors, and many other safety-critical functions (Rockwell Collins, 1999). Operating the FMS involves programming and inputting data, locating and verifying information, updating data, detecting errors and inconsistencies, and problem solving, all of which is carried out in a high pressure situation with tight time constraints, and multiple demands and interruptions. Failure to program the FMS quickly and accurately can result in incidents such as a planes veering into unprotected airspace, taking off at speeds too slow for the plane's weight, or nearly running out of fuel.

Figure 2. FMS Interface

Safety reports we reviewed emphasized the need for pilots to automate their FMS programming skills as much as possible; however, the FMS is not an easy system to master, as it has many display modes, an outdated user interface, and outputs that are not conducive to helping pilots visualize the results of their programming inputs. Thus, student pilots, and even experienced pilots, often avoid the self-study practice sessions needed to become highly proficient at FMS operation. In our guided analysis of the ASU training environment, we found FMS operations to be a weak technical area for some of its graduates.

Deficient procedural skill and technical knowledge of the FMS negatively impact other areas of flight training as well. We know from first-hand observation that students waste valuable – and expensive – flight simulator time as instructors have to stop the simulation to provide them with remedial FMS instruction. Students do receive instruction in FMS principles and procedures, and have access to a desktop computer simulation program for practice, but there is little opportunity to practice using the device in a structured, goal-oriented environment.

From this analysis, and what we have learned from our initial TARGET research, we concluded that development of an FMS serious game could improve FMS training efficiency and effectiveness based, in part, on the following criteria:

- ✓ Repeated practice is necessary to master the skills and knowledge required to execute the task quickly and accurately.
- ✓ There is a current lack of opportunity, and dedicated resources, for students to practice the task
- ✓ Performing the task is not inherently motivating.
- ✓ To be meaningful, the task should be carried out within the context of a structured, goal-oriented scenario (i.e., "free exploration" using a simulated FMS was likely to be ineffective).
- ✓ There is no clear, immediate feedback available during self-study using available simulators (i.e., unless an instructor is there to point it out, it is not always obvious to students whether an error was made, or how to fix it).
- ✓ Poor performance in this area can have significant consequences that impact other aspects of training.
- ✓ Students are not reaching or maintaining desired masterly level using conventional training methods.

Selection of Game Elements and Design Patterns

After selecting FMS operations as a good candidate for GBT, we then began to link relevant environmental parameters to best-fit game elements variation in order to develop the game. Because our TARGET database was still in an early stage of development, selection of these elements was primarily based on our research on game design and general instructional design principles.

Figure 3 depicts the rationale behind the selection of specific game element variations. For example, as shown the first row of the figure, we knew that students had some difficulty with integrating and performing component tasks, so we designed the game to have a whole-task environment with support from a virtual captain. We also noted that some of the component tasks – such as learning the key strokes to get to a specific FMS page, or being able to quickly locate data – would require repeated practice that might not be provided in a solely whole-task environment, so we also incorporated part-task activities to address this, as indicated in the second row of the figure.

We also looked at characteristics of the learner, as illustrated in row 3. The students who would be interacting with the game were generally high achievers and we observed a friendly competition among them; thus, competition is made salient by posting the scores of the top three players, as were rewards for exceptional performance.

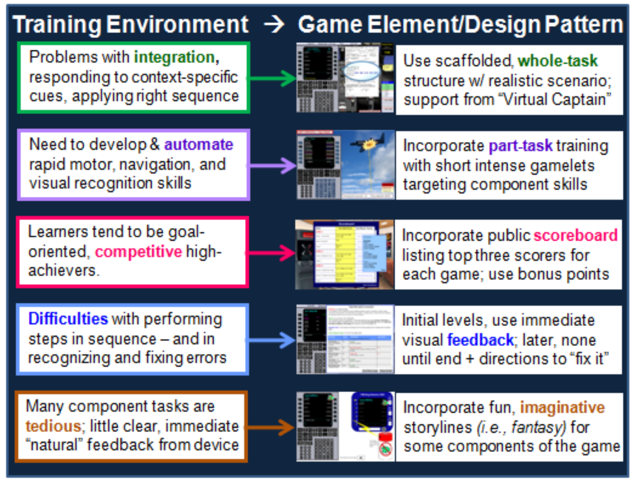

Figure 3: Linking Training Environment Parameters to Best-Fit Game Elements

As illustrated in row 4, another consideration was that students often had trouble adhering to the correct sequence of steps. Thus, in earlier game levels, the correct sequence is reinforced through visual feedback provided in an "evaluation checklist." Another important skill, however, is learning to verify one's own performance, by catching and correcting any errors that may have been made. Thus, in later levels, feedback is withheld until the end, where students must go back and fix any mistakes they may have made.

FMS Game Design

The outcome of this analysis was the development of an FMS game that comprises three main integrated components: (1) a context-rich whole task training environment that we called "Flights;" (2) focused, just-in-time part-task training in the form of mini arcade games; and (3) a central interface dubbed "The Pilot's Lounge" where players could easily access scores as well as the other two component of the game. (See Figure 4)

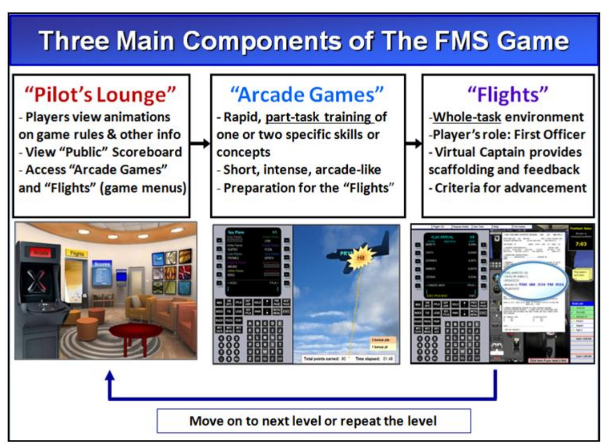

Figure 4: General Structure of the FMS Game

In the ***Pilot's Lounge*** (left panel of Fig. 4), players can: click on a "computer" to watch an animation explaining the rules of the game or view demonstrations of pilots carrying out specific tasks; view their scores and compare them to those of the top three players; and access the "Arcade" games and the "Flight" board.

Arcade Games (center panel of Fig. 4***)*** are short, rapid, part-task activities designed to familiarize students with component skills. In level 1 of the game, for example, this includes retro-type games such as "Spy Plane" and "Bug Kill" which provide practice with learning the FMS keyboard layout and scratchpad interface of the device – and help players build a mental model of how to use the keys and menus to navigate to specific pages. Each level of the FMS game has one or two such arcade games that players can access prior to moving on to the "flight" for that level.

The ***Flights*** are the "whole-task" component of the game, and have more cognitive and physical realism than the arcade games. Here, players must program the FMS device within a certain time window, while the game's Virtual Captain provides hints and feedback. During initial levels, the Captain provides explicit assistance to guide novice players through the sequence (e.g., "Go to the Perf Init page"). In later levels, assistance is only provided if the players ask for hints.

It is important to note that, although players move on to levels of increasing difficulty, they also have the opportunity to go back at any point in the game and repeat any of the arcade games and Flights from the previous levels.

FMS Game: Implementation and Testing

One of the main goals of the project was to gather clean empirical data on the effectiveness of GBT compared to more conventional methods of training. The problem with drawing useful conclusions from many of the existing game-based training research comparing GBT to more conventional methods of training is that there often was not a fair comparison between the two. If the game group ended up performing better, it was unclear if it was mainly due to the fact that they had received more information or had more opportunity to practice (Hays, 2005). We wanted a fair comparison, so in addition to creating the FMS game, we also created a conventional, computer based (CBT) version of the training, which had the same content information, the same number of exercises, and addressed the same skills – but was presented in an "exercise" format rather than a game format. Figure 5 illustrates some of the similarities differences between the two versions.

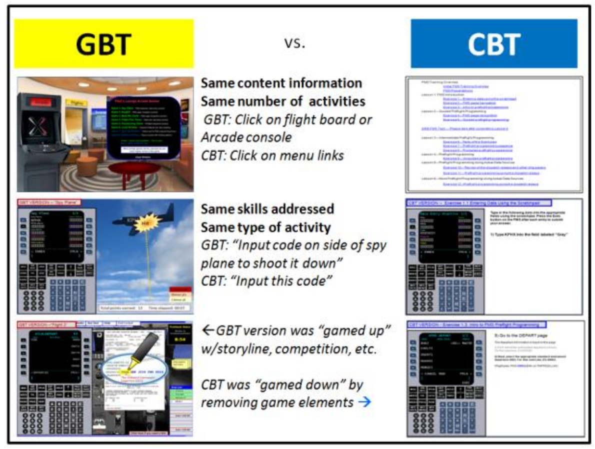

Figure 5: Comparison of Game-Based Training (GBT) & Conventional Computer-Based Training (CBT) Versions of the FMS Training Programs

General Procedure:

In all three studies, the students, all of whom were in their final year of a 4-year pilot training program to become commercial pilots, were randomly assigned to receive either the GBT version or the CBT version of the FMS training. All students were told they would be interacting with a new program designed to train basic FMS procedures and that they could play the games (or, for the CBT group, do the exercises) as many times as they wished. They were also told they could leave at any time, as long as they interacted with each game (or exercise) at least once. Students primarily interacted with the online training programs from the Department's lab computers, but could also access it from home if they wanted. Lab times were arranged so that only students in the same group were in the lab at the same time. They were also instructed not to discuss the program with students in the other group.

At the end of their respective training periods, students in both groups then took a criterion transfer performance test, where they were asked to carry out a preflight programming task using a handset FMS device and a commercial, virtual flightdeck desktop software program, (as shown in Figure 6).

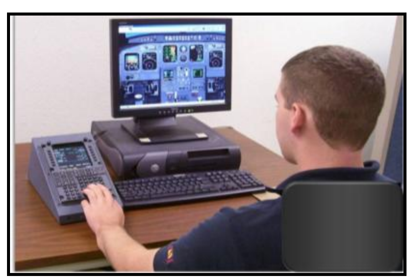

Figure 6: FMS Handset and Simulated FMS Program Used in FMS Criterion Transfer Testing

The test was administered to each student individually by an experimenter who was blind to the condition to which the participants had been assigned. The test sessions were videotaped, and scored based on accuracy, efficiency (minimal unnecessary moves), and time to complete the task.

GBT vs. CBT: FMS Study Results

Study 1: Introductory Training

In order to develop initial proof of concept, in the first study we gave students (n = 14) just the initial level of the training, which they had the opportunity to interact with during two one-hour lab sessions. One week after completing the training, they then took the criterion transfer test. As illustrated in Figure 7, we found that students who had interacted with the GBT version scored significantly higher on the accuracy measure, *making fewer programming and sequencing errors* than students who had interacted with the CBT version of the program *($p < .05$).* Students in the GBT group also completed the FMS preflight programming more quickly than students in the CBT group, though that difference failed to reach statistical significance.

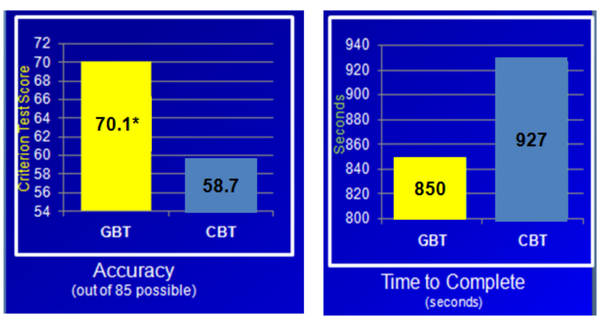

Figure 7: Study 1 (FMS Training Level 1) - Transfer Test Scores (* *$p < .05$*)

Study 2a: Intermediate Training, Levels 1-3

In Study 2, another set of students (n = 19), none of whom were in the first study, interacted with Levels 1-3 of the FMS training programs. All three levels provided repeated practice with basic preflight programming, using different scenarios (e.g., different flight plans and flight data values), with less direct guidance provided in the latter levels. One of the main questions addressed in Study 2 was whether the pattern of results observed in Study 1 would continue after both groups received additional training. In other words, does GBT offer just an initial performance boost and would the CBT group soon catch up? Levels 2 and 3 of the game also incorporated a new goal structure. After debriefing students from Study 1, we realized that the time goal (to complete the entire programming in a given amount of time) was not salient enough; we hypothesizes that it might be more effective to set smaller goals and have students earn bonus points if they accurately completed, for example, the first three sequencing steps within two minutes, the next set of steps in three minutes, and so on.

Ten students were randomly assigned to receive the GBT program and 11 the CBT program; however, two students were later dropped from the GBT group for technical reasons, leaving 8 in the GBT group. The students completed the training in three one-hour lab sessions spread out over two weeks. A week after completing the training, all students took the criterion transfer test (which, given the additional training students received, had been recalibrated to be worth 100 points). Again, as depicted in Figure 8, students in the GBT group made significantly ***fewer errors*** *($p < .05$)* than students in the CBT group. This time, however, the GBT students also ***completed the task more quickly*** *($p < .05$)*) than students in the CBT group.

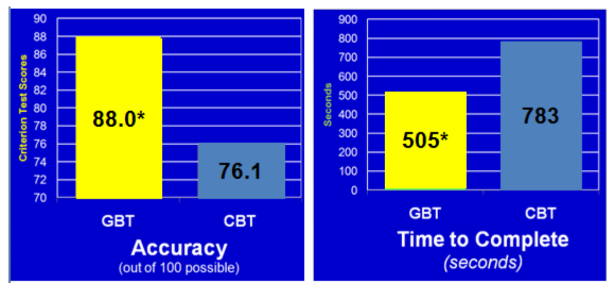

Figure 8: Study 2a (FMS Training Levels 1-3) - Transfer Test Scores (* *$p < .05$*)

Consistent with findings from the first study, we also found that students in the GBT group ***voluntarily repeated*** the training more frequently than those in the CBT group *($p < .001$).* Adding up the total number of times each student played the 8 games that comprise Levels 1-3 (or in the case of the CBT, 8 exercises), we found that students in the GBT group voluntarily interacted with the games an average of 61.5 times, while those in the CBT group only did the exercises 21.5 times. Again, the GBT games and the CBT exercises involved the same number of "problems" and entailed the same type of activity (typing in data, locating specific pages, etc.) – and in both conditions, end-of-game messages asked students if they would like to repeat the game (or exercise).

Study 2b: Delayed Retention

With Study 2, we were also interested in whether the benefits of GBT would be retained over time. To test this, one month after completing the FMS training – during which the students were not given access to the FMS games or exercises – we tested the students again. This time, rather than the FMS handset device, we used an online version of the FMS similar to the ones used in the "Flight" games and exercises, but with all hints and game elements stripped away. Again, as shown in Figure 9, students in the GBT group outperformed the CBT group on measures of accuracy and speed, thus suggesting that the benefits of GBT are not fragile.

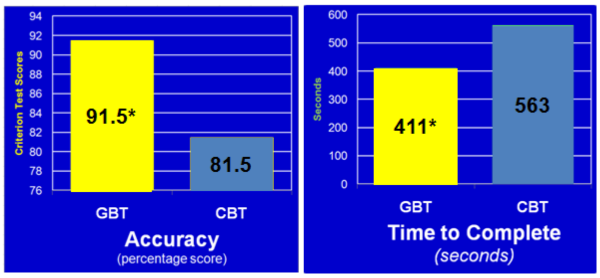

Figure 9: Study 2b (FMS Training Levels 1-3) – Delayed Retention Test Scores *(* p < .05)*

Study 3: Levels 1-6 of the FMS Game

In the third study, 14 students were randomly assigned to the GBT condition and 14 to the CBT condition; none of the students had participated in the earlier studies. The students in both groups interacted with the first six levels of their respective FMS training program over a two week period. The latter levels (levels 4-6) included more advanced FMS programming skills and required students to locate data points using actual data source materials (e..g., dispatch releases, cargo load slips, and electronic displays). The later levels also included new rules and a new feedback design pattern. In these levels, players were allowed to make errors, do steps out of sequence, and make inefficient moves without receiving immediate feedback. This allowed them the opportunity to go back to verify their work, recognize cues that signal errors (such as a discontinuity in the flight path), and then go back and fix them.

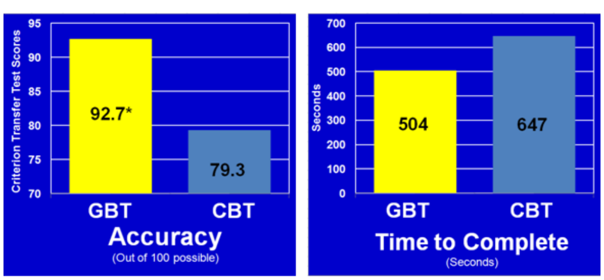

Figure 10: Study 3 (FMS Training Levels 1-6) – Transfer Test Scores *(*p < .01)*

A week after completing the training, all students took the criterion transfer test. As depicted in Figure 10, students in the GBT group made significantly ***fewer errors*** *(p = .009)* than students in the CBT group, and also ***completed the task more quickly*** than students in the CBT group, though this difference was only marginally significant (p = .07)

Anecdotal Evidence of Long Term Impact

By the time we ran the third study, some of the students who had participated our initial studies were now taking advanced simulator training with instructors from Mesa Airlines which partners with ASU. We heard, anecdotally, that some of these instructors – who were unaware of the new FMS training – had asked the ASU instructors if they were "doing anything differently with the FMS training." They had noticed that some of the current students were much more proficient with the FMS than previous students had been. As more students move into this advanced training, we will begin gathering empirical evidence of this longer-term impact of GBT on overall training.

APPLICATION TO AIRCREW TRAINING – PART 2: CREW COORDINATION TRAINING

After developing the FMS Game, our next step was to take what we had learned and, again, apply the TARGET methodology to identify another area within ASU's flight training program, that would likely benefit from the insertion of GBT technology. This time we focused on training gaps in the area of Crew Coordination. Using the criteria set forth in the TARGET model, we selected Profile-Based Callouts (PBC) as our targeted training area.

In the sections below, we provide a brief synopsis of the methodology used to develop and evaluate the PBC training game, including results of the empirical studies comparing the effect of the PBC game to conventional training on criterion test performance.

Analysis of the PBC Training Environment

Carrying out profile-based callouts is a crew communication task that requires pilots to use standardized procedures and terminology during critical phases of flight, and is an integral aspect of the safe and efficient operation of a multi-pilot aircraft. Part procedure and part technique, executing PBCs requires pilots to detect the particular profile situation the aircraft is in (e.g., missed approach, normal landing, avoidance maneuver), and then initiate the appropriate series of specific task actions and scripted verbal exchanges in coordination with fellow crew members. The level of accuracy and synchronization demanded by this task requires precise teamwork and fluid, seamless integration of physical actions and verbal responses. Given the number of different profile and callout sequences to learn – and the number of details embedded in each sequence – it is not an easy task to master. In fact, one of the instructors, a retired Southwest pilot, informed us that 80% of the first officers who fail their first year, "Off Probation" simulator proficiency check, do so for not knowing their "Profiles and Call Outs."

After conducting our guided analysis of the PBC training environment, we selected PBC as a good candidate for GBT based, in part, on the following:

- ✓ Repeated practice required for mastery.
- ✓ No dedicated program (e.g., computer program) available to train this task.
- ✓ Current training with one-on-one guidance from an instructor is generally effective, but is time consuming.
- ✓ Limited opportunity for full, whole-task practice (Task requires interaction between two crew members; difficult to practice on one's own).
- ✓ Requires precise attention to details and exact responses; without an instructor present, trainees might not notice errors.
- ✓ Tasks require practice using multiple scenarios – including unexpected events.

Selection of Game Elements and Design Patterns

Because we were able to use information gathered during the FMS Game design process, it took us significantly less time and effort to create the specifications and initial levels of Profile and Callout Game.

The PBC game is made up of several modules, each addressing a specific PBC sequence (e.g., Normal Takeoff, Precision Approach, etc.) Each module has six levels ranging from a quiz-type game in level 1, to a simulated cockpit environment the later levels (see Figure 11), where players respond to realistic cues. In some of the levels, we used modified design patterns from the FMS game to address skills shared by the two areas. For example, for one of the PBC games that focused on sequencing, we reused design patterns from one of the FMS's part-task "Sequencing Game."

Figure 11: Snapshot from Level 4 of the Profile and Callout Training Game

After we developed the first four levels of the PBC game, we then ran a study examining the effects of GBT compared to conventional training methods on PBC performance.

PBC Game: Implementation and Testing

As with the FMS studies, we randomly assigned students enrolled in their fourth year of ASU's Professional Flight Training Program to either the GBT or the CBT group. Students in the GBT group interacted with four levels of the "King Air, Normal Flaps 8 Takeoff" game module during two 30-minute lab periods, and also had the option of accessing the game from home. For the conventional group, we did something a bit different from what we had done in the FMS studies. This time, rather than create a separate computer-based "non-game" version of the training, we had students in the "CBT" group engage in *true* conventional training – that is, what they would normally do as new hires with the airlines, which is to study paper diagrams and notes of the profile/callout sequences (such as that shown in Figure 12) and then practice the sequence in a static simulator, without additional guidance or feedback from an instructor. Students in both groups, however, still received the same basic content information and had the same amount of time and opportunity to practice.

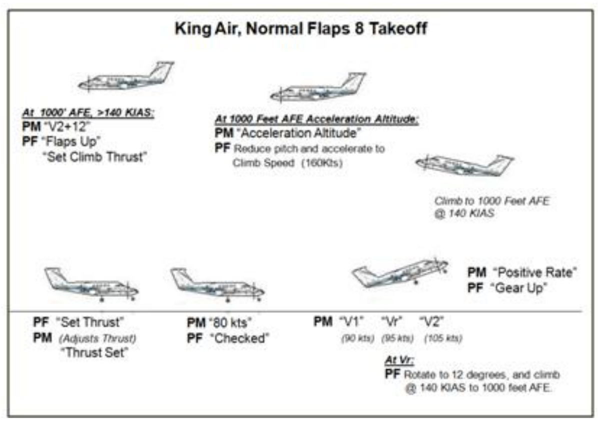

Figure 12: Diagram of King Air Normal Flaps 8 Takeoff Profile and Callout Sequence Used in the PBC Training

After students completed their respective PBC training with the first module (Normal Flaps 8 Takeoff), they then took the PBC criterion transfer test. During the test, students sat in the King Air Simulator and performed the profile and callout sequence for a Normal Flaps 8 Takeoff, assuming the role of the "Pilot Monitoring." Our trained confederate sat in the other seat and assumed the role of the Captain (see Figure 13). Students were scored based on how accurate they were in carrying out the callouts and actions, and on whether they performed them at the correct time and in the correct sequence.

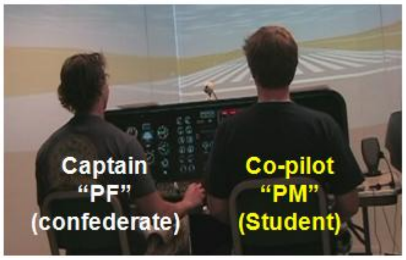

Figure 13: Set up for Profile and Callout Criterion Transfer Test with King Air Simulator

GBT vs. CBT: Profile-Based Callouts Study Results

As shown in Figure 14, analysis of the results indicated that those who received the GBT version of the training performed significantly better on the criterion transfer assessment than those who had received the conventional training, making fewer wording, sequence, and timing errors – thus providing further support for the TARGET model and for the efficacy of GBT.

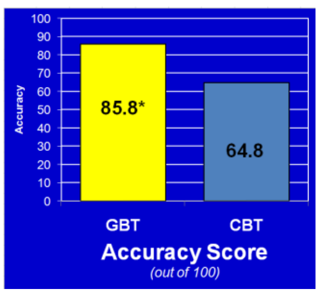

Figure 14: Results from Profile-Based Callout Study Criterion Transfer Test *(*p < .05)*

DISCUSSION

Based on our experiences in developing games to support FMS and PBC training, we are in a position to address some of the major questions that arise regarding the desirability of GBT. Firstly, while GBT may be "fun," is it actually effective? That is, does it add anything over and above CBT and will the acquired skills transfer to the simulator and field environment?

Based on the results of three empirical studies of FMS operation and our initial PBC study, the answer is a most definitive "Yes." That is, we found that students receiving GBT in FMS operation performed more accurately and faster in a criterion transfer task compared to a well-controlled CBT group. These results reached statistical significance in each study. Moreover, in a delayed transfer test, we found that GBT students still enjoyed a performance advantage over their CBT counterparts, indicating that the skills acquired through GBT are not inherently fragile and will hold up over a reasonable retention period. We also found a statistical advantage of GBT over current training for a very different task, the crew coordination-intensive profile-based callouts.

While our empirical results support the effectiveness of GBT, they do not give the instructional designer cart blanche to use *any* type of game to support training. Rather, the game elements must be carefully adapted to the specific training environment. We found that our decisions were greatly aided by conducting focus groups with the students after each study. These sessions informed us about the student reactions to published scores, positioning of timers, placement of feedback, and the relative uses of part-task vs. whole-task training.

Secondly, even if we accept that GBT can, under certain conditions, impact training effectiveness, *why* is it effective? Answering this question requires the conduct of a carefully designed series of research studies, which will consider, for example, such psychological processes as immersion and "flow." However, even from our initial studies, one likely mechanism is the amount of self-initiated practice that GBT induces vs. conventional training. In particular, we found in our study of the FMS that GBT students, on average, spent almost 300% more time practicing with the FMS than our CBT controls. This is, of course, one of the results we hoped to find with GBT – that students would voluntarily spend more time interacting with the material. As to why this is so, a plausible explanation is the expectations of gameplay, in which it is *expected* that one will repeat the game to beat others' scores or to get a better score. This expectation is not present for conventional training, where completing a textbook lesson or CBT exercise is the natural end point. To determine the relative importance of competition as a defining element of gaming, studies are needed that compare the impact of GBT with and without posted scores, where the frequency, timing, and level of detail in the scoring are systematically varied.

Thirdly, there is a concern that using a GBT approach is not advisable since "games" might trivialize training. Again, our studies did not find any evidence of such an attitude. That is, during the conduct of our focus groups, we found that the ASU students were able to differentiate between the more game-like part task games and the more sim-like whole task game. They understood that the "arcade games" were designed to aid performance on the "flights" and that the flights helped them understand and perform the actual task better. Thus, the students and instructors were quite accepting of the learning intent of GBT, though it also helped that we made the objectives and purpose of each game clear in the design of the overall GBT system.

Fourth, another common concern about the viability of GBT as a training tool is that it has to have high fidelity – and involve use of expensive games – in order to be effective. On the contrary, our FMS and PBC games were developed using Flash and simple graphics, none of which involved an excessive outlay in materials or development time. However, we believe that one way to reduce needless expense, and avoid risk, is to employ a science-based methodology to ensure that only those elements that have proven effective in a given training environment are developed. It is for this reason that we are continuing to enhance TARGET so that GBT design is as efficient as possible.

Fifth, there is a concern that GBT just adds another demand on limited instruction time and over-burdened instructors. However, GBT can be implemented with a relatively small time investment AND it can encourage more productive self-study. Since the GBT is deployed as a self-contained unit, it actually off-loads the instructor and, besides, having better prepared students results in more effective use of subsequent simulator and class time.

Finally, will GBT be an effective medium for training higher-order cognitive skills, as these are becoming an increasingly more prevalent aspect of military and commercial industrial tasks? In this regard, we are beginning to construct a series of web-based serious games that address a core set of cognitive meta-competencies that we believe are needed to support warfighters engaged in Irregular Warfare operations (Spiker & Johnston, 2010). Examples of essential meta-competencies include (a) switching between narrow and wide field of views to maintain situation awareness, (b) adopting someone else's perspective (vs. having an ego-centric view), and (c) recognizing an unfolding event by detecting a piece of it. Figure 15 shows an example screen from the web-based GBT we are developing to promote acquisition of the field of view switching meta-skill. We believe that providing GBT in these core meta-skills have enormous potential for achieving high returns on investment by covering multiple task domains with a single application.

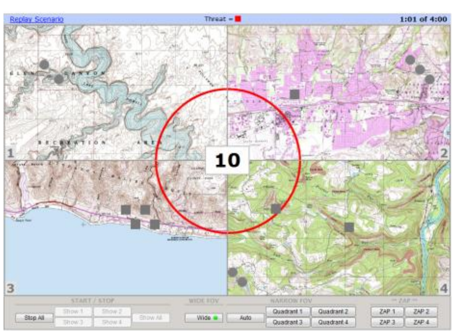

Figure 15: Snapshot from the Situational Awareness Training Game

REFERENCES

Aldrich, C. (2005). *Learning by doing: A comprehensive guide to simulations, computer games, and pedagogy in e-learning and other educational experiences*. Pfeiffer: San Francisco, CA

Bergeron, B. (2006). *Developing serious games*. Hingham, MA: Charles River Media, Inc.

Conkey, C. (2009). Machinima technology: Serious games lite. *Proceedings of the 29^{th} Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Garris, R., Ahlers. R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation and Gaming*, *33*, 441-467.

Hays, R. T. (2005). The *effectiveness of instructional games: A literature review and discussion*. Technical Report. Naval Air Warfare Center: Orlando, FL.

Mautone, P. D., Spiker, V. A., Dick, R. A. (2007). Design Patterns for Serious Games in Tactical Problem Solving. Paper presented at the 2007 International Conference on Software Engineering Research and Practice (SERP'07). Las Vegas, NV,

Mautone, P.D., Spiker, V.A., & Karp, M.R. (2008). Using serious game technology to improve aircrew training. *Proceedings of the 28^{th} Interservice/Industry Training Systems and Education Conference*. Orlando, FL

Mayer, R. E. (2003). Learning and instruction. Upper Saddle River, NJ: Prentice Hall.

Prensky, M. (2001) *Digital game-based learning*. McGraw-Hill: New York, NY

Rockwell Collins. (1999). *Collins FMS-4200 Flight management system: Pilot's guide*. (2^{nd} ed). Cedar Rapids, IA: Rockwell Collins.

Salas, E., Wilson, K.A., Priest, H.A., & Guthrie, J.W. (2006). Design, delivery, and evaluation of training systems. In G. Salvendy (Ed.*), Handbook of human factors and ergonomics* (3^{rd} ed), p 472-512. New York: John Wiley & Sons.

Spiker, V. A. (2006) Principles of Instructional Design to Support Scenario-based Training, Technical Report. Naval Air Warfare Center: Orlando, FL.

Spiker, V.A. & Johnston, J.H. (2010). *Using behavioral science principles to train small units*. First Cross-cultural decision-making conference. Miami, OH.