Using Serious Game Technology to Improve Aircrew 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

ABSTRACT

Serious games offer a promising technology for training complex skills. But there are few implementation guidelines and even fewer empirical studies that unequivocally demonstrate a benefit of game-based training over conventional alternatives. To address the first issue, we developed a structured query framework that links elements of game design (e.g., feedback, challenge, fantasy) to training objectives. The framework is a synthesis of serious game, multimedia instruction, cognitive psychology, and instructional design literatures. This tool helps instructors incorporate gaming elements into their curriculum to improve the motivation, learning, and performance of key cognitive, psychomotor, and problem-solving skills. This framework also specifies the types of gaming environments that work best for each skill. To address the second issue, we applied the framework to develop a serious game to train an operational task: programming an aircraft's flight management system (FMS). A training needs analysis revealed the FMS to be a good candidate for "gaming up" since its 1970's interface and opaque logic are unfamiliar to many of today's pilot trainees, where the intrinsic motivation to practice component skills is low. Yet learning to program the device quickly and accurately – and being able to recognize and correct errors – requires repeated practice in varied contexts. The training effectiveness of the FMS game is being evaluated using students from Arizona State University's flight training program and co-located Mesa Airlines' new-hires as subjects. A randomized design is used where half the students receive gamed FMS instruction and half conventional computer based training. A transfer of training task criterion test is then given using a physical replica of the FMS device, where comparisons of relative performance index game impact. The paper will provide graphic examples of the framework and quantitative results of the evaluation.

ABOUT THE AUTHORS

Tricia Mautone is a senior scientist at Anacapa Sciences, Inc. and has a Ph.D. in cognitive psychology from the University of California, Santa Barbara. Her work primarily focuses on the development and evaluation of effective multimedia and game-based instruction. Since joining Anacapa Sciences in 2003, Dr. Mautone has worked on several projects involving computer-based instruction and testing, such as designing the structure and format of interactive crowd control training modules, creating and evaluating a web-based training program targeting the development of critical thinking skills in military personnel, and developing and applying principles of cognitive psychology to the development of serious games for science and technology instruction.

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. Dr. Spiker specializes in the analysis of user functional requirements, design of advanced information training systems, and evaluation of human cognitive performance. Dr. Spiker received the Human Factors Society's Alexander C. Williams Award (1991) for outstanding human factors contributions to the design of a major operational system.

Merrill R. (Ron) Karp is Professor of Aviation and Associate Department Chair, Department of Aeronautical Management Technology, Arizona State University, Polytechnic Campus. Since 1994, he has developed and managed the ASU/Mesa Airlines Airline Bridge Training Program using airline high-fidelity motion-based and fixed-based simulators and scenario-based training. Dr. Karp received his Ph.D. in Administration & Management from Walden University in 1996. He was a graduate from the U.S. National War College in 1985 and the Armed Forces Staff College in 1981. 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 served three combat tours in the F-4 during the War in Vietnam.

Using Serious Game Technology to Improve Aircrew 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

INTRODUCTION

Using games in education has a long history, but the use of serious games – games specifically designed for education and training – is a natural response to challenges faced by civilian and military organizations in such areas as reduced manning, trainee attrition, and the growing need for cross training. Coupled with less time and fewer resources for field training, there is an increasing demand for innovative approaches that maximize training efficiency and effectiveness (McDowell, 2007). Serious games as a new training technology would appear to have many advantages, as recent advances in computer graphics have reduced the time and cost to produce high-end computer games (McDowell, 2007; Smith, 2006) while many of today's student-trainees have been brought up playing computer games. It is no wonder, then, that organizations now look to incorporating serious games into their training curricula, as game based training (GBT), to reduce training costs and time and to increase learner motivation and training effectiveness (Gee, 2005; Hays, 2005; Leemkuil, de Jong, & Ootes, 2000).

However, GBT is no panacea, as (1) there are few empirically-based guidelines on when and how to implement GBT; (2) there are virtually no cleanly designed empirical studies in which the performance of GBT-trained students was compared to conventional (as opposed to no) training; and (3) it may be that fewer than half of today's students actually play games with any regularity (Belanich, et. al, 2007). This paper describes an ONR-funded [1] project whose primary objective is to conduct such a study, using an actual training enterprise as our "laboratory." The audience for our study was the Arizona State University (ASU) aircrew training program, where the goal was to use GBT to improve training in operating the aircraft's flight management system (FMS). To design the serious game properly, we need a tool that links specific training objectives and learning environments with specific, best-fit game elements. The tool we are developing for this purpose, TARGET – the Tool for Applying Robust Gaming Elements to Training - indicates whether GBT is a good candidate for a particular training situation, and offers recommendations for optimal game element variations to embed in instruction. Before describing this tool and our study, we first describe the elements of serious games that give rise to GBT.

USING SERIOUS GAMES FOR TRAINING

Like "regular" games, serious games present goals, rules and constraints, interactivity, discovery, feedback, challenges, and competition within a storyline involving role-playing and social interaction, and thus are designed to engage and motivate players. But unlike regular games, serious games also have defined learning objectives, real-world relevance, and incorporate integrated instructional support such as prescriptive feedback, scaffolding, and progressive levels of difficulty. Looking at the elements that comprise a serious game, its utility as an instructional tool is understandable. First, many game activities are remarkably similar to the *instructional objectives* of training programs. A well-designed game may require a player to: problem-solve; develop fundamental skills; use tools effectively; synthesize information; learn and apply rules; plan and strategize; visualize outcomes; induce connections between events; or understand the risks of making decisions (Prensky, 2001) – all desired skills in almost any domain. Second, serious game-playing factors seem to *parallel good pedagogy*. Games are based on interactivity and active participation, where to do well, players must be cognitively active. Games also often offer ample opportunity for practice and repetition, which is essential for automatic skill acquisition (Anderson, 2000). Providing immediate, clear feedback is another common game element that mirrors good pedagogy, as is the graduated complexity built into many games that allows players to progress at their own pace. Finally, games often entail situated learning – role-playing in a realistic, context-rich environment – allowing practice in situations impractical to replicate in field training (Mayer, 2003). But the most-often cited advantage of using game elements in instruction is *that games are motivating* – people *want* to play games and often devote a great deal of time and focus to mastering the skills and rules

[1] The research described in this report was funded by STTR contract no.: 68335-07-C-0386. The technical point of contact is Curtis Conkey.

of the game (Aldrich, 2005). A premise behind GBT is that incorporating game elements into training will increase student engagement and time on task, thus improving learning outcomes and perhaps learners' attitudes toward the subject matter as well.

While it makes sense to use games for training, does incorporating serious game elements into training actually improve performance? Unfortunately, existing research is sparse and fragmented (Hays, 2005). Some studies show that games do result in improved learning outcomes, (e.g., Gremmen & Potters, 1995; Garris, et al. 2002) while others indicate that games had little positive impact, as players focused more on gaming aspects than on the intended learning objective (Hays, 2005; Reiber & Noah, 1997). A problem in interpretation is that many of the studies did not specify the gaming elements employed, nor even what learning outcomes were measured (Hays, 2005). Rather than asking "Are serious games effective?" we should ask, "How can we best implement GBT to maximize performance and motivation?"

To address these questions, we must first develop a common framework to describe serious games. A comprehensive taxonomy of serious game elements would allow us to more precisely and consistently define serious game elements and to catalog their many variations. Having this common language would aid in generalizing research findings and allow us to develop research-based principles linking specific characteristics of a target training environment to best-fit candidate game elements.

THE SERIOUS GAME ELEMENT TAXONOMY

We constructed a preliminary taxonomy of serious game elements to identify, categorize and more precisely define some of the more salient serious game elements and their variations. The initial content and structure of the taxonomy was derived by extracting serious game elements from a review of the literature on games and serious games, as well as literature on instructional design, multimedia design, and cognitive and educational psychology. We drew upon existing lists of game dimensions (e.g., Garris, et al. 2002, Hays, 2005; Leemkuil, de Jong, & Ootes, 2000), combining some dimensions, such as curiosity, mystery, discovery, and novelty, into one category (Curiosity and Discovery), and adding other categories – such as Feedback and Instructional Support – that were missing from some lists, but were often cited in reviews of educational games as being critical factors influencing the effectiveness of instructional games (Hays, 2005; Leemkuil, de Jong, & Ootes, 2000). Figure 1 shows the ten elements that comprise the taxonomy. To save space, we restrict discussion to the five elements that were most useful in developing the FMS training game.

Five Main Elements

1) ***Feedback and Scoring*** refers to information given to the player about progress toward a goal. It can vary from a simple confirmation that a particular action or choice was correct, to elaborate, prescriptive information about what the player did correctly and incorrectly, and what changes are needed to improve performance. Different training situations may require different feedback variations. For example, in training rapid response skills, providing quantitative, event-specific feedback is more appropriate than delayed or elaborate feedback, which may disrupt the player's flow of responses (Hays, 2005).

2) Rules and Constraints refer to the boundaries of the game and the consequences for adhering to or violating these boundaries. They encompass factors such as instructions for play, range of allowable actions, rewards and penalties, and general path of play. Rules and constraints affect both learning outcomes and

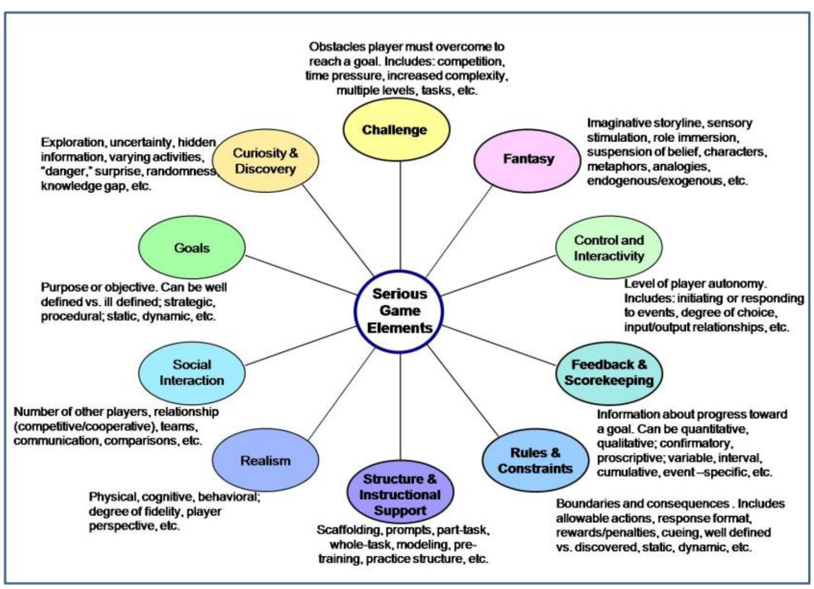

Figure1: Serious Game Elements Taxonomy

player motivation. For example, repeatedly penalizing a player for failing the same task, without providing any opportunity to pursue a different path, practice the requisite skill, or obtain remedial help, will likely result in the player quitting the game (Leemkuil, et al. 2000).

3) Challenge refers to the obstacles the player must overcome to reach a goal. Injecting challenges in a serious game is particularly important when the skills targeted by the game are inherently dull. To engage players and encourage them to practice mastery of these skills, the game must build in motivating challenges. Competition – with other players, with a virtual character, or with oneself (e.g., beating a time) – is a common way to introduce challenge. Incorporating increasing complexity, time stressors, twists, and uncertainty also increases the challenge (Aldrich, 2005).

4) Structure and Instructional Support refer to the scaffold, such as hints, pre-training, and focus questions, that help learners understand how to use and learn from the game, as well as other key instructional factors such as how practice sessions are structured. While this category would seem at first to mainly affect learning outcomes, it is also a critical factor in affecting player motivation. In any medium, learners are more likely to be motivated when they can make sense of the material (Kintch, 1980; Mayer, 2003).

5) Fantasy emphasizes gaming elements relating to imaginative storylines, role-playing, sensory stimulation, and suspension of belief. This may be especially relevant if the underlying component training tasks are not inherently interesting and, thus, where motivation needs bolstering. On the other hand, some players, particularly adults who are training for professional development, could be turned off by what they perceive as a frivolous or "childish" storyline.

Using the Taxonomy to Guide Game Design

The taxonomy helped us organize the preliminary findings from our review of the serious game literature, particularly findings from empirical studies examining the impact of instructional games on learning outcomes. It also allowed us to address the key questions: When are serious game elements effective, and how are they best implemented?

Consider the taxonomy category "*Instructional Support and Structure*." Two ways to design the structure of a serious game are to employ (1) a whole-task approach, where players encounter the full game environment at all times, or (2) a part-task structure, where they master subtasks before encountering the full game environment. Does the way a game is structured affect learning outcomes? If so, under what circumstances is one structure more appropriate than another?

Using our taxonomy, we looked for studies that touched upon that issue. For example, Fabiani, et. al (1989) compared the effects of two different game structures – a part-task structure and a scaffolded, whole-task structure – in a game designed to train fighter pilots to handle the cognitive and attentional demands during flight. When task demands – number of interruptions – increased, those who received the whole-task version outperformed part-task subjects. Thus, it appears that while a part-task game structure might be appropriate in some situations, when the targeted training environment requires learners to handle increasing task demands and interruptions, a scaffolded, whole-task structure may be more effective.

Although existing research on the effectiveness of games is rather sparse and poorly described, we were able to use the taxonomy to distill salient game features and their impact on learning in specific training environments. We then used this data to develop a crosswalk linking specific game element variations to specific training environments. Our goal is to transform the crosswalk into an XML relational database on effective game design principles. Using the serious game taxonomy as a guide, we are populating the database with findings from the serious game literature, and from research in cognitive psychology, educational psychology, and instructional design. As described below, we are adding data from our own empirical studies of the effects of different game element variations in different training environments. The relational database will form a cornerstone of TARGET and will aid instructional designers in determining when and how to implement game-based training.

OVERVIEW OF TARGET

Still in development, TARGET will ultimately consist of three main components: a query tool, an XML relational database linking game elements and training environment characteristics, and a set of recommended game element configurations or design patterns. The ***query tool*** will scaffold the training systems analysis and aid instructional designers in creating a profile of a particular training environment. It will identify characteristics of the environment (e.g., type of learners, tasks, cognitive functions required, typical problems encountered with training) relevant for selecting appropriate game elements. The ***relational database*** component will match details of the training environment profile with game element variations (e.g., method of feedback, level of fantasy, type of challenges, etc.) most appropriate for that environment. This match-up will, again, be based in part on findings from cognitive psychology and serious game research. The output will be a ***set of recommended game elements*** that entail a written description (e.g., "recommend using a scaffolded, whole-task approach" followed by a definition, example, and rationale for the approach). The output may also link to an item in a reusable design pattern library (e.g., a concrete example that can be incorporated into another game, along with associated implementation specifications), or even a link to a sample game, accompanied by suggestions for modifications. Or, the output might suggest that a game is not the most appropriate method for instruction in this situation. This could happen 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.

We have begun applying principles from TARGET (i.e., creating a profile of the training program, determining if it's a good candidate for GBT, matching characteristics of the training environment to game elements, generating and testing recommended game element configurations) to create a game to target a specific training area. Information gleaned from this process will be used to refine the TARGET tool specifications.

THE FMS GAME

The competency selected for our current project is training student pilots how to program the flight management system (FMS) on a Canadair regional jet (CRJ) aircraft. We selected this task from a training needs analysis of the aircrew training program at ASU, our research partner institution. Below, we briefly describe FMS operations and why it is an appropriate area for GBT. We then describe how we developed the FMS programming game using the serious games taxonomy and the TARGET methodology.

Overview of FMS Operations

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, set up of multi-function displays, radio tuning, fuel-efficient flight, and many other safety-critical functions (Rockwell Collins, 1999). The FMS (Fig. 2) is a complex electronic device whose operation involves programming and inputting data, locating and verifying information, updating data, detecting errors and inconsistencies, and problem solving, all carried out in a high pressure situation with tight time constraints, multiple demands and interruptions, unpredictable events, and serious consequences for failure. 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 looked at 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 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 ongoing research with ASU, we have 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 not much opportunity to practice using the device in a structured, goal-oriented environment. In addition, the task itself is not particularly motivating, so students don't put in the required practice, which, of course, has negative consequences later on. 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.

Development of the FMS Game

Once we determined that GBT was appropriate for the FMS environment, we began developing training objectives and specifications for a game designed to target those objectives. From our task analysis of the FMS, and discussions with subject matter experts at ASU, we distilled the main knowledge and skills that student pilots must master in order to operate the FMS efficiently, as well as some of the problems typically encountered when learning to use the FMS. These problems include: failing to follow the appropriate sequence, entering data in the wrong fields, entering incorrect data, failing to enter necessary information, not recognizing the current display "mode," failing to detect errors and discontinuities, and not entering the required data quickly enough in time for takeoff.

Because TARGET was still in development, we used the Serious Game Element Taxonomy and the game element-training environment crosswalk to select the element variations to incorporate into our game. Space limits preclude detailing all the selected game elements, so we will highlight a few. Table 1 provides a summary

Game Element Taxonomy Category	**FMS Training Environment Characteristic**	**Recommended Game Element Variation**	**Example of Application to FMS Game**	**Rationale (supporting study)**	***Counter Example***
Structure & Instructional Support #1	Students need big picture view; need to juggle tasks, handle interruptions; typical problems include integrating tasks	Use scaffolded, whole- task structure	Incorporate cockpit environment; have players program full preflight sequence; virtual captain provides scaffolding	When learning to handle multiple demands, whole-task structure is effective (Fabiani, et al., 1989)	*You may not need this for training environments with low multitasking or task integration problems*
Structure & Instructional Support #2	Students need to develop rapid motor & visual recognition skills; typical problems include slow data input	Have separate part-task structure with short, intense activities	Include Pilots' Lounge Arcade: Short, intense gamelets that target component skills	Focused, repeated practice is best for automatizing motor skills (Anderson, 2000)	*You may not need this with high ability students who just need task integration practice*
Realism	Students need to apply skills and procedures in a specific context; not always safe or practical to practice skills in actual environment	Incorporate high degree of visual, behavioral, and cognitive realism	Interface should look and function like real device; storylines and tasks fairly realistic, employ increasing stressors	High degree of fidelity fosters transfer to domain-specific environments (Mayer, 2003)	*May not need added cost of high physical realism if training problem is primarily handling cognitive demands*
Fantasy	Component tasks (e.g., data entry, page navigation) are "boring"	For training of some component tasks – create fun, amusing storylines	Pilots lounge arcade games have high level of fantasy (cracking safes, killing monsters) with time constraints	Imaginative storylines can increase learners' motivation and task persistence (Malone, 1981)	*May not need a high degree of fantasy if motivation not an issue (e.g. .training flight control)*
Feedback	Many tasks consist of several steps and subtasks	Use branched prompts and feedback to shape behavior; guide successive approximation	Virtual captain first provides subtle prompts, then successively more explicit hints per player request	Complexity of feedback should match level and type of learning (Spiker, 2006)	*Not needed for more holistic tasks, such as performing scripted call outs*

Table 1: Summary of Rationale for Selection of Game Elements for and FMS Training Game

of our selection rationale, beginning in the first column with the Serious Game Element Taxonomy category from which they were drawn. The second column lists characteristics of the FMS training environment relevant for selecting particular game elements; the third column describes the variation of the game element that our crosswalk indicated would best fit these training environment needs; and the fourth column provides an example of how the game element variation was applied in the FMS game. The fifth column provides a rationale for why the game element was selected, in the form of a research finding or general principle the literature. Finally, column six provides a counter example of a training situation in which the selected game element variation might not be appropriate.

The second row addresses the taxonomy category, *"Structure and Instructional Support."* In programming the FMS, students need to have a "big picture" understanding of how to interact with the FMS in a real-world, cockpit environment. This includes following the appropriate sequence while programming the FMS in preparation for takeoff, as well as handling the time pressures of performing the task with frequent interruptions. Therefore, we elected to structure the game using a scaffolded, whole-task approach where students are exposed to the whole task throughout the training, but receive guidance and prompts (from a "virtual captain") prior to carrying out some of the steps. This way, students receive repeated experiences with the entire process – and learn how to juggle multiple task demands, as well as handle interruptions – but are not cognitively overwhelmed. Research in both general instructional design (Mayer, 2003; Sweller 1999) and GBT (Fabiani, et. al 1989) suggest that this type of scaffolding and cognitive apprenticeship helps learners understand the steps and see how the whole process works, and also better prepares students for the cognitive demands of the actual training environment.

On the other hand, many of the component FMS tasks, such as using the appropriate buttons and line select keys to rapidly navigate to specific pages within the FMS, involve procedural motor and visual skills. Developing these skills requires a great deal of repetitive practice before students can reach a level where performance is automated (Anderson, 2000). To provide the type of practice needed, we chose to include a set of rapid, part-task training tasks – in the form of arcade-type games – that players can access between levels of the whole-task game. As summarized in the third row of the table, these games provide players with the repetitive practice needed to automate the component skills. Our crosswalk also indicated that a part-task game structure is particularly effective with beginning learners, our intended population.

The fourth and fifth rows of the table address Realism and Fantasy. Programming the FMS is something that trainees will be carrying out not only in the ASU simulators, but very shortly, in their professional careers as pilots. Because the skills and procedures students learn during training will be applied directly to actual work environments, it is important that students have the opportunity to experience how members of their profession think, behave, and solve problems. This is referred to in the cognitive literature as distributed authentic professionalism and situated learning (Aldrich, 2005; Gee, 2005; Mayer, 2003). Therefore, for the main game, we created a simple but visually and aurally realistic interface that replicates much of the functional experience of a cockpit-resident FMS. The game also includes elements of "fantasy" or suspended realism. In the real world, a novice would not be given the responsibilities that players in this game are given, nor would they be able to see the consequences of their actions (e.g., seeing the plane veer off course, or even crash). In addition, to encourage players to practice some of the more "boring" tasks, the game also incorporates a set of arcade-type where players hone component skills, such as FMS page navigation and data recognition skills while attempting to "kill" monster-like bugs that are destroying FMS data.

An example of one of the different types of *prompts and feedback* employed in the game is highlighted in the last row. At the beginning levels of the game, players learn to carry out a series of steps involved in programming the FMS preflight. This involves learning the proper sequence of steps, the page they need to go to in the FMS to carry out that step, the buttons they need to press to get to that page, and what to do once they are there. Because the task consists of a series of subtasks, and because students have varying (but limited) prior experience using the FMS, we used a branching feedback and prompting structure, where players may request successively more explicit guidance, as needed. This ensures that less experienced students do get too far off track, and that more experienced players can progress at their own speed.

FMS Gameplay

Below we summarize the main features of the FMS programming game, illustrating how the game elements and pedagogical principles were implemented.

The FMS game is divided into two parts: the "Flights" and the "Pilot's Lounge." The **"Flights"** provide whole-task training and takes place in the simulated cockpit of a CRJ Jet, as shown in Figure 3.

Figure 3: View of Game Cockpit, Level 1

The game provides players with a context-rich, realistic environment where they assume the role of the first officer who must program the FMS preflight sequence, update the FMS as new information arises, and detect errors and solve problems, all under the guidance of a "virtual captain" who provides prompts and feedback. The interface includes: a game control toolbar that provides help and lets players repeat audio, view audio as text, view scores, and set other options; a realistic interactive device with an underlying database that corresponds to real-world function, and which can be expanded (as shown here) or reduced by clicking on it; a countdown timer and a task stressing display (the luggage being loaded) that corresponds to actual time constraints on the flight line; and supporting materials, such as ship papers and maps. Players who successfully complete a flight within the allocated time and with minimal errors, have the option to repeat that flight for a better score or to progress to the next flight where they take on more responsibility.

The second part of the FMS game is the **"Pilot's Lounge,"** where players can go between "flights" to reinforce skills they will need in order to complete flight on subsequent levels. Here, they can play short, intense arcade-like "gamelets" that reinforce (via part-task training) the harder-to-master component skills from the flight portion of the FMS game. Figure 4 shows two of the gamelets. The top panel depicts “Spy Plane”, where players are familiarized with the FMS data entry interface, including the keyboard, line select keys, and scratchpad, as they type in codes painted on the sides of circling spy planes as rapidly and as accurately as possible. The bottom panel shows a snapshot from “Bug Kill”, where players must use their knowledge of FMS organization to navigate to pages and “kill a bug” before it destroys the data on that page.

In addition to the mini arcade, the pilot’s lounge also contains a scoreboard where players can view their own best scores and those of the top three players. The Pilot’s Lounge also has two communication devices, a cell phone and computer, from which players receive calls from virtual “captains” and “fellow pilots” regarding game play instructions and further information about FMS training.

Figure 4: Two Pilot’s Lounge Games

EXPERIMENTAL STUDY

To determine whether the FMS game leads to better training outcomes compared to a more conventional training program, and to gather data on the effect of the selected game elements described above, we conducted a controlled study examining the impact of the game elements on a training outcome performance assessment. In this, the first of a series of planned studies, one group of pilot trainees interacted with the FMS game while another group (the conventional-based training or CBT control group) received a “gamed down” or conventional, computer-based version of the FMS training. This latter group had the same information and practice opportunities as those provided in the game, but with the game elements (competition, fantasy, etc.) removed. We then compared how the two groups performed on a criterion task where trainees were asked to program the preflight sequence for a hypothetical flight. This controlled design allows us to draw conclusions about the effectiveness game-based training without the confounds often found in other studies.

Method

Participants: Fourteen subjects (all male) participated in the study. All were seniors currently enrolled in ASU Aeronautical Management Technology flight training program’s Airline Instrument Procedure class. Half were randomly assigned to the game-based version of the training program and half to the conventional version. The average age of all participants was 23.2 years (SD 2.5) and they reported an average of 3.8 (SD 3.8) hours of prior experience using an FMS (excluding one student in the CBT group who reported over 100 hours of prior FMS experience). In the weeks prior to the study, all students had viewed lecture slides describing the features of the FMS, and had had a few hours of practice inputting data into a simulated FMS.

Materials: The training materials for both the GBT and the CBT groups were created using Flash 9.0 and were presented to participants via Internet Explorer. The programs were designed so that non-game factors, such as instructional content and number and general type of exercises, did not vary between the two programs. The CBT version could thus be considered nearly identical to the GBT version except that game elements (storyline, characters, competition, explicit challenges, etc) were removed and the game tasks were presented in a conventional exercise format.

Training materials for the GBT group consisted of Level 1 of the FMS game described above. It began with an introduction to the game environment – including the pilot’s lounge, public scoreboard, cockpit features, and the arcade. After orientation, players were instructed by the “virtual captain” to play the two arcades games, “Spy Plane” and “Bug Kill,” to prepare them for “Flight 1”. For both gamelets, point totals and reaction time scores were displayed both during and at the end of the game. In the Flight 1 game, players assumed the role of first officer and had to program the FMS preflight sequence (gate to pushback) with explicit audio instructions from the virtual captain about what page to go to and what steps to carry out. The captain provided hints whenever a player hit the “hint” button, and also gave the player feedback (e.g., “Good!” or “Not quite, try again.”).

Training materials for the CBT group consisted of the same number and type of exercises as those in GBT, but with game element features removed. Figure 5 shows a snapshot of Exercise 1.1 (the CBT equivalent of Spy Plane) where users are given exercises such as “Type KPHX into the field labeled gray”.

In Exercise 1.2 (the CBT equivalent of Bug Kill), players were given instructions such as "Go to the STATUS page," and then used the onscreen FMS function buttons to navigate to that page. The tasks and prompt in Exercise 1.3 were similar to those in Flight 1, but the players did not assume a role, interact with

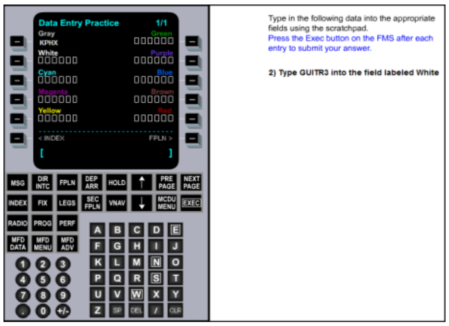

Figure 5: CBT Training, Exercise 1.1

virtual characters, view a cockpit interface, nor see the timer. For all three exercises, a textbox with the words "Correct" or "Incorrect, try again" appeared after each task item, and at the end of the exercise, the total time to complete the exercise was displayed.

Procedure: For the training part of the study, participants were run in groups of seven, with each participant seated in front of a computer equipped with headphones. Those in the CBT group were run in one 2-hour session, while those in the GBT group were run in a second 2-hour session. Participants 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. One week after the training, participants were given a criterion transfer test where they were required to complete an FMS preflight programming sequence for a simulated flight from Phoenix to San Diego. They were run one at a time in the ASU Airline Research Computer Lab, where they and the experimenter sat in front of a computer set up with a simulated FMS program and the FMS handset device (see Figure 6).

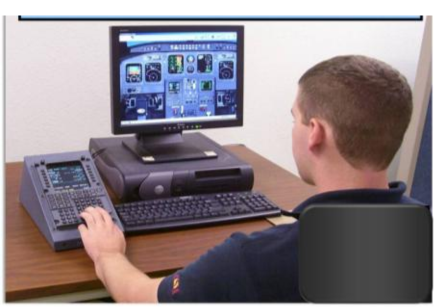

Figure 6. FMS Handset & Simulated FMS Program

Participants were told to go through the preflight programming steps as quickly and as accurately as possible, in sequence, starting with the Status page and ending with the Radio page, using the data listed on a provided data table. They were also told they could ask for a hint, but that the experimenter would not volunteer any information unless directly asked.

The experimenter then videotaped the screen of the FMS handset to record player's actions and comments. If participants indicated they were finished, but had actually skipped a step or two, the experimenter had them go back and complete that step. When all participants had been tested, the experimenter conducted a group debrief to gather specific feedback regarding the game and conventional training programs. During the debrief, all participants were shown slides of both the CBT and GBT programs. They were asked general questions, such as what they liked and didn't like about the game, as well as more focused questions about specific game element features, using the serious games taxonomy as a guide.

Results

Scoring. The videotapes were scored by an experimenter who was blind to the condition to which the participants had been assigned. Scoring was done using a point system rubric that assigned points for completing a programming step in the correct sequence, as well as points for completeness and accuracy of each step. Points were deducted for skipping steps, inputting data in the wrong fields, inputting the wrong data, asking for hints, and navigating to irrelevant screens. The total number of points possible was 85. Participants also received a time score which was calculated from the time they began the first step to the time they completed the last step – including the time when they were instructed to go back and complete a step they had skipped.

Effect of GBT on Criterion Test Performance. Table 2 shows the means and standard errors for criterion test performance for the GBT and CBT groups.

Table 2. Criterion Transfer Test Performance

	Accuracy Score (points)		Time to Complete (seconds)	
Group	**Mean**	**(SEM)**	**Mean**	**(SEM)**
GBT	70.1	(2.3)	850	(86)
CBT	59.7	(3.8)	927	(108)

Results of an independent sample t-test indicated that participants who received the game-based version of the FMS training scored significantly higher on the accuracy measure than those who had received the conventional version of the training, $t(12) = -2.32$, $p = .04$. However, there was no significant difference between the two groups in terms of how long it took participants to complete the test $t(12) = .56$, $p = .59$. This finding suggests that those who completed the game-based training were more accurate and more complete in carrying out the FMS preflight programming in the proper sequence, but were not necessarily faster. This may be because some of the more accurate participants took time to verify and check their entries, while some of the less accurate ones did not.

Effect of GBT on Motivation. After completing each game or exercise, participants in the GBT and CBT groups, respectively, were given the option of repeating them as many times as they wished. It was predicted that those in the GBT group would find the games engaging and challenging, and thus opt for more repeat practice than those in the CBT group. According to the results of an independent samples t-test, this prediction was borne out, $t(12) = -2.30$, $p = .04$. After playing each of the three games one time, as instructed, those in the GBT group elected to play one or more of the games an average of 2.33 more times (SD = 2.50), while those in the CBT group repeated one or more of the exercises an average 0.14 times (SD = 0.37). In fact, while five of seven GBT participants repeated one or more of the games, only one out of the seven CBT participants did an exercise more than once.

Discussion

One of the main goals of this study was to establish whether trainees who received GBT – in particular, GBT designed using the principles outlined in the taxonomy crosswalk described above – exhibited better training outcomes compared to those who received conventional or “non-game” training. Based on participants’ performance on the criterion transfer test, in which they carried out a realistic FMS programming task, it appears that the additional practice that students received while interacting with the FMS game – even just level 1 – led to improved FMS programming skills over those who received more conventional training.

A second objective of the study was to glean information about the relative effectiveness of specific game elements based on observations of the players’ behavior, as well as from feedback during the debrief. Using the taxonomy to guide our evaluation let us make more focused observations and ask specific questions in order to draw conclusions about the impact of specific game elements and how they did (or did not) contribute to the effectiveness of the game. This, in turn, allowed us to generate specific solutions for improving the game.

Below, we summarize observations regarding the impact of integrating three serious game elements into the FMS game. As well, we discuss: how the element was implemented; the impact of that element variation; and implications for how we will use that information to not only improve the current FMS game, but also to refine our TARGET database and thus aid in the design of other games for training areas that share features with the FMS training environment.

Feedback and Scoring Element Variations*:* For the arcade games, average reaction time scores and total points were displayed after each player response. However, participants reported that they did not attend to the immediate feedback, as there was too much information to absorb in a limited time amount of time. They did, however, notice their final end-of-game scores and were motivated to improve them. On the other hand, in the flight part of the game, students reported that there was too little information about how they were doing, overall, other than how much time remained, and that they didn’t really have an overall sense of how well they were doing.

Implications: When speed *and* accuracy are important, simplify the scoring and use cues such as color coding to make the feedback more salient and more easily interpreted. For longer (whole task) games, provide more salient and continuous feedback of overall accuracy and efficiency, in addition to immediate feedback after each response. Based on these observations, for the updated version of the FMS game, we plan to simplify the scoring of the arcade games and to incorporate a meter representing the virtual captain’s level of approval, which goes up or down depending on the number of errors and unnecessary actions a player makes.

Rules and Constraints: For Spy Plane, we wanted to discourage guessing and encourage accurate data entry. We thus included a “rule/constraints” variation whereby players were given a limited number of “lives” and the game ended after a player made three errors. Based on observation of play patterns and on player feedback, however, we realized that rather than encouraging the players to be more accurate, this rule just led to frustration and confusion about what they had done wrong.

Implications: For introductory-level training of a fast-paced skill, impose penalty points for errors rather than deducting “lives” -- at least for the first level. In the next version of the FMS game, we will eliminate this rule and instead impose more salient penalty points for errors and continue to display the errors until players correct them.

Realism Elements and Fantasy Elements: In this first study, we deliberately included games with varying levels of fantasy. Flight 1 was fairly realistic with only a little fantasy in the form of role playing whereas Spy Plane had a fairly simple storyline, and Bug Kill had a high degree of extraneous fantasy. The general feedback we received was that students felt there was no need for an elaborate storyline or the addition of fantasy elements that were not relevant to the training environment. For example students reported that they liked assuming the role of first officer and interacting with the virtual captain; however, answering phone calls or navigating through the simulated airport was distracting. Unlike with an entertainment game, they did not need to have a fully immersive storyline; instead, they preferred more direct challenges (such as "complete this before the clock runs out" or "meet these accuracy criteria") as well as more direct competition. They also indicated that they thought the storyline behind the Bug Kill game was not really suited for people training to be professionals, though they said that the Spy Plane game, which had a more straightforward storyline, was fine.

Implications: For adult learners who are highly competitive, embed games either in a realistic context that they find motivating (e.g., assuming the role of professional pilots) or issue direct challenges they can see are explicitly relevant to their training needs.

FUTURE DIRECTIONS

In subsequent studies we will attempt to replicate and extend the above finding by using a larger sample of pilot trainees and by incorporating additional levels of the FMS game into the training experience. Likewise, we will also integrate more complex FMS operations into the criterion test, such as programming holds and route changes, to see how the effect of GBT on those problem-solving tasks. In order to provide more empirical support for the guidelines specified in our TARGET tool, and to test the hypotheses generated from the findings of the first study, in these future studies we will use a variation of the ablation method, where we remove a particular design feature (e.g., time stressors) from the game, or swap it out with an alternative variation (e.g., confirmatory feedback rather than branched feedback) and then compare the resulting learning outcomes. This will allow us to further analyze the contribution of specific game features to the learning experience. We will also develop plans to create a series of game design patterns that will allow us to transfer successful components of the FMS game to other domains, such as operating the Navy's amphibious landing craft air cushioned (LCAC) vehicle, machine operation, and other related domains (Mautone, Spiker, & Dick, 2007), where we will conduct a third set of studies examining the games' effect on training outcomes in different domains.

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

Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed). New York: Worth

Belanich, J. , Orvis, K. A., Moore, J. C., Horn, D. B., Solberg, J. L. (2007). Fact or Fiction – Soldiers are Gamers: Potential Effects on Training. Paper presented at the 2007 Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) Orlando FL

Fabiani, M., Buckley, J., Gratton, G., Coles, M. G. H., Donchin, E., & Logie, R. (1989) The training of complex task performance. Acta Psychologica, 71, 1-3.

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

Gee, J. P. (2005). What would a state of the art instructional video game look like? retrieved October 2006/ www.innovateonline.info/index.php?view=article&id=80

Gremmen, H. & Potters, H. (1995). *Assessing the efficacy of gaming in economics education*. Retrieved March 15, 2006 from http://arno.uvt.no/show.cgi?fid=3278.

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

Leemkuil, H., de Jong, T., & Ootes, S. (2000). *Review of educational use of games and simulations*. Project Report Knowledge management Interactive Training System consortium: Ae Enschede, The Netherlands

Malone, T. W. (1981) Heuristics for designing enjoyable user interfaces: Lessons from computer games. *Association fro Computing Machinery*

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, June 2007

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

McDowell, P. (2007). Serious Games: Why Today? Where Tomorrow? MS&T Magazine, Issue 2/2007.

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

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

Smith, R. (2006). Game Impact Theory: The five forces that are driving the adoption of game technologies within multiple established industries. Retrieved May 2006 from (http://www.ctonet.org/)

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

Sweller, J. (1999) Instructional design in technical area. Camberwell, Australia. ACER Press.