Helicopter Flight Training Through Serious Aviation Gaming

Michael D. Proctor
Department of Industrial Engineering and Management Systems
University of Central Florida
mproctor@mail.ucf.edu

Maria Bauer
US Army RDECOM Simulation and Training Technology Center

Thomas Lucario
Aviation Officer, US Army Simulation Operations

The notion of Serious Games dates to at least 1970. More recently leaders also identified gaming technology as a possible disruptive technology. If off-the-shelf PC-based aviation games and the vast library of related civilian developed databases and models can be leveraged for serious training use, then existing flight training paradigms from familiarization training to mission rehearsal might be disrupted and provide the military financial, safety, quality, and time benefits for even less cost. This research investigates the contribution that an off-the-shelf, PC-based, aviation game makes to learning using three inexpensive interface configurations. The simulator performance improvement methodology is used to measure the contribution. The research considers interface usability, model fidelity, and simulation sufficiency for task learning. The research also investigates the difference in performance of pilots with and without turbulence with increased load in these configurations. The specific task chosen for research was combat search and rescue with turbulent weather conditions. All forty-five participants in the research were in training to become licensed helicopter pilots. Results of their subjective assessments are also included.

Keywords: aviation, training, personal computers

1. Introduction

The notion of Serious Games goes back to at least Abt in 1970 [1]. More recently Smith identified computer games as technology that could potentially disrupt current industrial paradigms in the manner described by Christensen [2, 3]. Off-the-shelf, Personal Computer (PC)-based, aviation games may be posed as a potential disruptive technology in the area of aviation flight simulation training.

A number of hurdles must be crossed before current paradigms of simulation flight training might be disrupted. Current flight simulators meet specific training requirements and comply with either FAA guidance for civilian applications or meet DoD acquisition requirements prescribed by the funding program [4, 5, 6]. Current flight simulators and training devices are composed of computer-based authoritatively verified and validated models and simulations of environmental parameters and aircraft flight dynamics [4, 5]. Further, they provide human interfaces that display some portion of the environment and may replica to some degree instrumentation panels, controls and physical motion of the aircraft [4, 5]. If off-the-shelf, PC-based, aviation games are to be truly considered a disruptive technology in the field of flight simulation then these games must credibly address the same training requirements of the current technology that they might displace.

The objective of this research is to advance the level of understanding of game technology as a technology disruption agent in general and the understanding of an off-the-shelf, PC-based commercial game as a disruptive force in aviation training in particular. The commercial game X-Plane is of interest in this research as the game provides a comprehensive set of subsonic and supersonic flight dynamics, scenery of the entire earth between -60 and +74 degrees of latitude, and variable weather conditions to include wind shear, turbulence, and micro-bursts [7]. Specific contributions of this research include insight into the application of this game to a specific training application but also insight into the near term potential for such games to create technology disruption in flight simulation training.

This article presents discussion, objectives, and measures of effectiveness about aviation game technology in the sections below. The section on viability considerations addresses cost, system availability, target audience, and general component requirements. The section on training transfer methodology presents the methodology used in this research to estimate the effectiveness of a training system. The interface and model consideration section identifies interface components and models topics that were a focus of the research. The section on training media describes the three media applications used in the research. The remaining sections address scenario, evaluation protocols, results, and conclusion.

2. Viability Considerations

Given the infancy of research into the use of game-based software on a PC as a system for helicopter training, the U.S. Army Research, Development, and Engineering Command sponsored research to determine the potential contribution of an off-the-shelf, game-based system for mission rehearsal aspects of helicopter pilot training. Specific strategic objectives included low upfront cost and the ability to deploy the system to remote areas.

A PC-based game or even simulation may not be an appropriate media to train many aviation tasks. For example, other than startup and shut down procedures, Stewart et al [8] indicate that for actual manipulation of the aircraft controls for flight “the U.S. Army does not use simulation in the primary (contact) phase of initial entry rotary-wing (IERW) training.” Further the capabilities and risk of PC-based games in aviation training is not fully understood. At the time of the previously cited research, the Army Research Institute was “in the process of undertaking a new series of transfer of training experiments, using a low-cost, PC-based simulator which represents the TH-67.”

Therefore games to be seriously considered for aviation training must provide a compelling or substantial motivation to a sponsor or set of sponsors to disrupt the current training paradigm. Motivation may involve any of the cost-benefits identified by Edward Link as far back as 1929 when his simulation technology disrupted the then current training paradigms [9]. Ever since then simulation has disrupted and replaced more costly training paradigms by achieving monetary and time savings, improving safety, increasing training system availability, etc. [9].

From a cost perspective, we sought an off-the-shelf, PC-based commercial game to meet one of our primary interests of low cost upfront funding. X-Plane met that requirement. If one can meet training needs with an off-the-shelf game, then the development costs are avoided by taking advantage of the huge investments made by the commercial developer of the game. The magnitude of the cost benefit ascribed to a game is also tied to the nature and size of the training audience and the scope of the tasks possible for training. Cost for simulation and/or gaming as a training tool vary across a considerable range. Factors impacting cost include whether one may simply purchase an off-the-shelf, PC-based game as we did in this research or whether one must create, design, develop, field, and support the game or simulator system. Other game alternatives exist besides off-the-shelf purchase such as proprietary PC-games built specifically for the training or non-proprietary PC-games built from open sources [10]. Proprietary and non-proprietary games on console-based and hand-held devices may be viable disruptive technologies in the future.

Another motivating interest in games on our part arises from the desire to train experienced military pilots in remote locations for possible mission rehearsal. This inferred low cost and easily transportable solutions that contained a vast readily available database. X-Plane and Microsoft Flight Simulator are two such off-the-shelf, PC-based commercial games that met that threshold consideration. For example, Laminar Research X-Plane contained at the time of this research 40 aircraft models and 18,000 airports across the United States and overseas [7]. This represents a vast database of models that might be useable for military purposes and thus avoid the expense of building and maintaining those databases. Further, X-Plane includes special effects such as day/night and wind and applies weather conditions typical of military operations through a flight model described as "blade element theory." This theory attempts to achieve realism by breaking the aircraft wings, horizontal stabilizer, vertical stabilizers, and propeller(s) if equipped into finite number of elements and then finding the aerodynamic forces on each element many times per second [7]. Victim and hostage rescue scenarios of interest to this research often require representation of severe air turbulence with the increased load of victims rescued. Finally, aerodynamic force representation is necessary to qualify as a "simulator" by the FAA [4, 5].

Finally given the wide number of different aircraft types, a game might have greater technology disruption potential if it functioned with many different interface systems. X-Plane can be configured to a number of different interfaces, which enable it to be used across aircraft types and training applications. As evidence of that, recently the FAA granted Fidelity Flight Simulation approval to train fixed wing pilots toward commercial certification, instrument rating, and the airline transport pilot certificate using a full motion simulator that uses X-Plane [11]. Recognition of X-Plane for serious training by this credible outside source contributed our choosing it for this research.

To make this research possible, we needed helicopter pilots. Inexpensive new pilot training under hazard conditions is an interest of Bristow Academy, formerly Helicopter Adventures Inc of Titusville, Florida. Bristow Academy trains prospective helicopter pilots from all over the world. Part of that training involves hazard landings of helicopters on off-shore oil platforms under turbulent air conditions. Since X-Plane provides serious gaming capabilities that include air turbulence and the ability to connect to multiple interfaces, Bristow Academy provided hanger space for the game-based simulators that we examined and access to pilots for this research.

3. Training Transfer Methodology

For a game to be considered for serious training applications, it must also demonstrate an acceptable level of training transfer for a given audience and task. Roessingh [12] and Talleur, et al [13] identify studies and uses as far back as 1990 that focus on applying PC-based games in initial flight training, but could not substantiate transfer of complex manual flying skills to the aircraft. Typically due to the lack of funds and safety hazards, evaluation of training transfer to a real aviation environment is beyond the scope of theoretical research. Instead the transfer of training methodology specified in "Simulator Performance Improvement Model" tends to be an acceptable substitute [14]. This methodology uses observed performance changes in the training system as measures of effectiveness of the simulator for training (Figure 1).

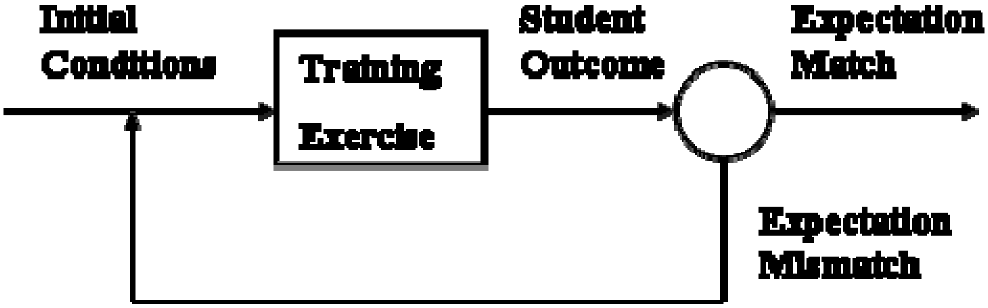

Figure 1: Model of Performance Evaluation and Improvement

Initial conditions include the demographics and experience level of the intended student audience, instructional delivery types and modes, hardware and software components, and task objectives, conditions, and standards. Specifically within this methodology, student learning and performance outcome differences between a baseline and performance run in the simulator or training device serves as evidence that training is effective. Research that uses the methodology shown above provides only "indirect evidence of simulator effectiveness" [15]. Real-world longitudinal research that arises on intended training audiences and tasks form the basis for final determination of effectiveness of a particular simulator configuration. Longitudinal research was beyond the scope of this research.

4. Interface and Model Considerations

Interface usability, model fidelity, and the suitability for training of scenarios composed within the simulator influence the overall capability of the system to achieve training goals. Assuming simulation through games is potentially desirable for the students and suitable for the tasks to be trained, the acceptability of the game interface as a serious substitute for actual aircraft controls and other physical aspects of the environment must be thoroughly considered. Past research indicates that many typical game interfaces do not replicate actual aircraft controls and performance environment. For example, Proctor et al [16] found that U.S. Army pilots were generally receptive to the use of personal computers equipped with Microsoft Flight Simulator for training; however, visual limitations created by the traditional single monitor display and joystick control of the gaming system were detrimental. Specifically, the “China hat” on the joystick was used to swivel the visual field of regard and did not adequately support intercockpit team situation awareness and team performance training in a simulated multi-ship helicopter operation.

When seriously considering the suitability of gaming interfaces as media, one must begin with understanding the intended platform used by the customer. From an Army perspective, helicopter flight simulation is of most immediate importance. In terms of the helicopter interface, perhaps the aspect of simulation that has been one of the most controversial topics is the level of contribution that the presence or absence of motion makes to the training of helicopter pilots [8, 17, 18]. Hays, Jacobs, Prince and Salas [19] summarize the motion literature by indicating, that while motion has contributed to some positive outcomes when results were analyzed on a task by task basis, helicopter transfer of training studies have yielded no significant differences in performance between the group trained with motion from that of the group trained without motion. Another area of controversy is the visual display. Visual display controversy is discussed further below but relates to both display area and field of regard.

Based on the funding, three inexpensive configurations were considered, which will be discussed further below. Other potential interface considerations not addressed due to research limitations include the other senses such as auditory and olfaction input and output.

5. Training Media

With game software determined, our next step was to identify interfaces to be used as media in the experiment. The interfaces needed to have the potential of addressing our sponsors’ requirements for low cost helicopter training under turbulent air conditions. As mentioned X-Plane software is able to support multiple media sets for training. We sought inexpensive sets that could not only be available in a hangar for pilot experiential and re-enforcement learning, but also be forward deployed and support mission rehearsal training. The first media set considered was a generic helicopter training device with motion platform (Figure 2). During the experiment it was used in two modes, motion on and motion off. The device consisted of a 2DoF electro-mechanical motion system (pitch: +43/-67, roll: +43/-43), the Cabin and the Display Case. The Cabin included two seats (pilot and co-pilot), two joysticks, a pilot collective and two sets of rudder pedals. The Display Case provided a 60" (diagonal) rear-projection, 1024 x 768 resolution visual display system and housed the computer that operates the system.

Figure 2. Generic Helicopter Training Device with Motion Platform (courtesy of *Simulation Entertainment Group, Inc.*)

The second set of training media was a commercially-available, commonly-used, PC-based desktop system that might be used as a desktop trainer (figure 3). This desktop system used the same computer and X-Plane software configuration, and functionally identical collective, joystick, chair and pedals as the first training media set. These interface components were employed directly out of the box in accordance with the manufacturer's instruction without additional modification. The computer monitor used in the desktop trainer was a 19" Dell Trinitron. The resolution of the monitor was set to be identical to the Display Case (1024 x 768). The brightness and contrast of the monitor were calibrated to be roughly equivalent to the 60" display.

Figure 3. Basic Rotor Wing Hardware Package (courtesy of *Flight Link Aviation Training Devices*)

X-Plane version 7.61 was used in all three configurations in order to provide a consistent Synthetic Natural Environment, instrumentation display, and aircraft flight model. No direct modifications to the code were made, though some input variables were modified for this research. The computer system consisted of: Intel Pentium 4 3.06 GHz CPU, 1GB RAM, ATI Radeon 9700 Pro AGP Graphics Card, 40 GB IDE Hard Drive with 8MB Cache, Windows XP Operating System and SoundBlaster Audigy2 soundcard .

Given the equivalence of the scenario, instrumentation display, aircraft flight model, Synthetic Natural Environment, functionally identical collective, joystick, chair and pedals, the configuration differences were the primary drivers for pilot learning. Our null hypothesis was that there was no difference in the learning of the pilots over the course of their training due to system configuration.

Given the desire to potentially deploy this equipment into forward operating areas, we also considered the impact of the two largest contributors to cost, size, and weight—the monitor and the motion system. As such we designed the experiment so that we could discuss the contribution of the monitor size to performance by comparing the performance of pilots trained using the 19" monitor of the desktop system with the performance of pilots trained using the 60" monitor in the Cabin with the motion turned off. To evaluate the contribution of motion to performance, we consider and discuss the performance of pilots trained using the Cabin with motion off with the performance of pilots trained using the Cabin with motion turned on.

6. Mission Rehearsal Scenario and Turbulence

The scenario we used is a modified part of a large international command and control research scenario generated by The Technical Cooperation Program [20, 21].

In this scenario, each pilot flew a simulated UH-60 aircraft on a flight plan that involved two segments. The first segment originated at home station and progressed to a remote location in mountainous terrain to rescue a down crew. After pickup, they then proceeded on the second leg to an away airport at a similar distance during which each pilot experienced considerable air turbulence. Performance of the aircraft on the second segment of the flight also changed due to the weight of the additional passengers.

The two legs of the final flight allowed us to evaluate pilot performance with and without turbulence and increased load in each of the three configurations. Helicopter control under conditions of air turbulence with increased load is a particular challenging area for training. The control task was made more difficult as the second segment of the two segment flight plan contained moderate air turbulence and micro bursts. Air turbulence also provided us the opportunity to gain insight into the influence of the presence or absence of motion on helicopter control. Though the flight dynamics of underlying UH-60 model contained in X-Plane was not tested against a manufacturer provided model, subject matter experts subjectively validated the X-Plane flight dynamics and interface controls prior to the experiment.

Our null hypothesis is that there was no difference of the performance of pilots in the absence of turbulence or the presence of turbulence with increased load. Measures of pilot performance focused on learning to fly along a prescribed corridor safely and in a timely manner. Some specific measures included time out (pilot became lost), aircraft crashes, and deviation in heading and altitude control beyond acceptable standards. Measures are discussed in more detail below.

7. Evaluation Protocols

The outline of our experiment consisted of administration of: (1)Background Questionnaire, (2) Immersive Tendencies Questionnaire, (3) Familiarization Training on the system (0.5 hour), (4) break, (5) Practice trials in the simulator (3 trials, first will be used as baseline), (6) break, (7) Combat Search And Rescue (CSAR) final mission run, and (8) Feedback Questionnaire. After device familiarization, no instructor intervention occurred during or between these runs. This supports assessment of the degree to which each interface configuration (without instructor intervention) might contribute to experiential and re-enforcement learning over the set of runs (improving/strengthening pilot performance through experience).

An Immersive Tendencies Questionnaire consisting of sixteen questions was given to determine each participant's pre-disposition toward simulation. During the familiarization phase, each pilot became comfortable with the controls, asked questions about the device, and experienced a scenario similar to the one used in the Combat Search And Rescue (CSAR) final mission run. The Feedback Questionnaire distributed at the end of the experiment obtained subjective remarks on effectiveness of the training system with respect to helicopter flight skills, value of the system in a flight school, and suggested improvements to the research study.

8. Assessing Learning and Performance

For the purposes of this experiment, learning was defined as a statistically significant (.05) improvement in task performance from the baseline run to the CSAR run. If the null hypothesis was not rejected, then no difference in helicopter control occurred and therefore no learning. If the null hypothesis was rejected, then a statistically significant difference in pilot learning occurred. The direction of the change indicated whether or not an improvement occurred.

Learning aircraft control in each training configuration was assessed from three perspectives:

(1) Analyzing, using the Wilcoxon Signed Ranks Test, how the pilot complied with four Go/No Go performance measures during the CSAR mission when compared against the baseline trial.

(2) Comparing, using the Chi-Square Test, the number of crashes and timeouts (pilots were "timed-out" if more than 8 minutes have passed after takeoff without arriving to the pickup zone) in the CSAR mission against observations during the baseline trial.

(3) Analyzing, using the Wilcoxon Signed Ranks test, how the pilot complied with the US Army standard [13] for heading, speed, altitude for level flight for each flight segment in the baseline run compared to the CSAR run.

The second null hypothesis tested the hypothesis that performance in the turbulent flight segment was equivalent to performance in the non-turbulent flight segment. Using the Wilcoxon Signed Ranks Test, the outputs of the CSAR mission during the Non-Turbulence segment were compared against the ones of the same mission during the Microburst/ Moderate Turbulence segment to determine the impact of turbulence on pilot's performance for each of the three training configurations.

The performance criteria were based on military subject matter experts and military references as found in related Joint and US Army publications [22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33].

Due to the lack of funds and safety hazards, evaluation of training transfer to a real environment was beyond the scope of this research. Instead we used the transfer of training methodology specified in "Simulator Performance Improvement Model" [14]. The methodology used shows only "indirect evidence of simulator effectiveness" [15]. With this methodology, learning and performance differences in the training device serves as evidence that training is effective.

9. Experimental Results

Forty five subjects participated in the study. All subjects were helicopter pilots. They were assigned to one of the training configurations Cabin with Motion, Cabin with No Motion, and Desktop. The three groups had the same number of beginner, intermediate and advanced level pilots. Results are provided below in three groups: Learning, Performance, and Subjective Assessment. The learning differences by system are described by twelve measures and are shown in Tables 1 through 5. The performance difference with and without turbulence with increased load are described by three measures and are shown in Table 6. The subjective feedback results are discussed and in part shown in Tables 7.

9.1 Learning

Learning from baseline run to final run: Helicopter Control GO/NO GO Results

Pilot's learning in a given system over the trials was partially measured with respect to compliance with the GO/NO GO performance measures and is summarized in Table 1.

The Wilcoxon Signed Ranks Test was used to compare the results of the baseline run with the observations during the CSAR mission.

Table 1. Experiential Learning from Baseline to Final Run: Four GO/NO GO Measures

Configuration	Arrived at pickup zone within 8 minutes	Arrived safely at the landing zone	Overall mission accom-plished within 20 minutes	Followed corridor
Cabin Motion	0.0273	0.0313	0.0313	0.0078
Cabin No Motion	0.0273	0.0002	0.0002	0.0137
Desktop	0.125	0.0625	0.0625	0.0313

At the .05 level of significance, there were statistical differences in the performance of participants and therefore learning in all categories for all configurations except for the Desktop configuration. The Desktop configuration demonstrated statistical significance only for the "followed corridor" measure.

Crash Avoidance and Mission Timeout Learning

Pilot's learning helicopter control with respect to avoiding crashes and timing was analyzed using the Chi-Square statistic. The number of crashes and timeouts in the baseline run was compared with the number of crashes and timeouts during the final run. Results are shown in Table 2 (pilots were "timed-out" if more than 8 minutes had passed after takeoff without arriving to the pickup zone).

Table 2. Learning from Initial Trial to CSAR Mission: Crashes and Timeouts

	Crashes/Timeouts	
Configuration	No Turbulence	Turbulence
Cabin Motion	0.0528	0.0271
Cabin No Motion	0.0067	<0.0001
Desktop	0.2636	0.2723

At the .05 level of significance, there were statistically significant improvements in the performance of participants and, therefore learning, in both the Motion and the No Motion configurations for the Turbulence segment and for the No Motion Cabin configuration in the Non Turbulence segment. The Desktop configuration did not demonstrate any improvement.

Learning from Baseline Run to Final Run: Control of Heading, Velocity and Altitude, Flight Segment Analysis

Pilot's learning from the baseline run to the final run was analyzed, using the Wilcoxon Signed Ranks Test, for each flight segment with respect to the US Army standard [22] for heading, speed, altitude for level flight. A Visual Basic macro was developed to calculate the amount of time during level flight (for both the Turbulence and the Non Turbulence segments) that the pilot was out of the established ranges for heading, speed and altitude. The results are summarized in Tables 3, 4, and 5 below.

Table 3. Learning Heading Control

Configuration	Heading Non Turbulence	Heading Turbulence
Cabin Motion	0.0039	0.0313
Cabin No Motion	0.0391	0.0020
Desktop	0.0625	0.0625

At the .05 level of significance, the null hypothesis of equivalence in Heading Control was rejected for both the Cabin with Motion and the Cabin with No-Motion configuration for both the turbulent and non-turbulent flight segments. Therefore, the alternate hypothesis of learning heading control is accepted in the Cabin with Motion and with No Motion in both the Non Turbulence and the Turbulence flight segments. The Desktop configuration did not reject the hypothesis of no learning in either condition.

Table 4. Learning Speed Control

Configuration	Speed Non Turbulence	Speed Turbulence
Cabin Motion	0.3028	0.1272
Cabin No Motion	0.2293	0.0040
Desktop	0.0906	0.3203

Table 5. Learning Altitude Control

Configuration	Alt. Non Turbulence	Altitude Turbulence
Cabin Motion	0.1514	0.6250
Cabin No Motion	0.0730	0.4238

Desktop	0.0054	0.4648

With respect to Speed and Altitude Control, the null hypothesis could not be rejected for all other configuration, turbulence, and control combinations except for two. During the Turbulence segment, learning to control speed occurred in the Cabin with No Motion configuration. The Desktop configuration supported learning in only one area—altitude control during the Non Turbulence segment.

9.2 Performance

Performance Différences: Non turbulence vs Turbulence

The null hypothesis of no difference between the non-turbulent and turbulent flight segment was evaluated using the Wilcoxon Signed Ranks Test. The outputs of the Non-Turbulence flight segment of the final run were statistically compared at the .05 level of significance against the Microburst/ Moderate Turbulence flight segment for the same configuration of that same final run. This analysis does not compare configurations directly but rather compares the ability of the pilot to control heading, velocity, and altitude in the non-turbulence segment with those same variables in the turbulent with increased weight segment within the CSAR final run.

The levels of statistical difference are shown in Table 6 below.

Table 6. Effects of Turbulence with Increased Weight on Pilot's Performance

Configuration	Heading	Velocity	Altitude
Cabin Motion	0.5000	0.0002	0.0001
Cabin No Motion	0.5000	0.0026	0.0001
Desktop	0.2500	0.0001	0.0001

The null hypothesis of equivalence was rejected for both velocity and altitude in all three simulator configurations. Hence the participant’s performance in the three training configurations was negatively affected by turbulence with increased weight with respect to the velocity and altitude parameters. This was not unexpected as control of the helicopter in turbulence was expected to be more difficult than under conditions other than turbulence. The null hypothesis of equivalence for heading could not be rejected in any of the three simulator configurations.

9.3 Subjective Assessment

Immersive Tendencies Questionnaire

An Immersive Tendencies Questionnaire was provided to all the participants at the beginning of the training program to identify the level of prior experience with immersive environments. Some of the beginner participants that received high scores on the questionnaire were more successful than intermediate or advanced participants in avoiding crashes and timeouts. However, none of the statistical tests performed relating the scores (total score, Focus, Involvement, Games) to the pilots' performance produced a statistically significant result.

Feedback Questionnaire

The feedback questionnaire was provided to all forty-five participants at the end of the training session. Table 7 below indicates levels of statistical significance between the null hypothesis of absence of a problem with the identified attribute. Blank cells indicate no responses in that category. Statistically significant differences in three cells are highlighted above in bold (lack of feedback in the controls of the Cabin with No motion, and lack of sufficient detail in the terrain display for both the Cabin with motion and the desktop).

Comments on the lack of feedback in the controls of the cabin with motion turned off is a dichotomy unexplained by the data, especially since participants in the same cabin with motion turned on indicated no such feedback concern. One possible explanation put forth by participants is that the expectation for motion, since this was a Cabin sitting on a motion platform yet no motion was provided, led to negative feedback on the controls.

While not statistically significant, the 100 milliseconds motion latency estimated by Simulation Entertainment Group, Inc. for responsiveness and pedal control for the Cabin with Motion may have contributed to those complaints, though this level of latency, according to McDaniel et al. [18] is consistent with most trainer standards.

Table 7. Comparison against Null Hypothesis of No Problem with Training System Attribute

	Cab Motion	Cab No Motion	Desktop
Needs motion		0.1071	
Controls			
Slow response	0.1052		
Lack of control		**0.0225**	0.2217
Pedals	0.1052		0.2217
Mounting			0.1052
Display			
-Terrain references	**0.0088**	0.2235	**0.0484**
-Peripheral vision	0.4631		0.1052
-Control panel	0.4631	0.2235	
Sound	0.4631		0.4631

What is more interesting is the concern about the lack of sufficient detail in the terrain display. Participants from the three training configurations (seven from the Motion, three from the No Motion and five from the Desktop) commented that the terrain needed higher fidelity ground features in order to be able to judge speed and distance. This is a concern that supports advocates of improved fidelity in the Synthetic Natural Environment models. Desire for increased terrain and feature data had the most widespread concern as can be seen in Table 7. The three configurations did not show a statistically significant difference from each other since they were all using the same terrain and feature database. The display of the terrain, however, differed by monitor size, and viewing of the terrain differed due to motion conditions.

10. Summary, Conclusions and Discussion

This research investigated an off-the-shelf, PC-based, aviation game for the capability to seriously model and simulate UH-60 flight dynamics with air turbulence and varying aircraft weight for the purpose of training. Specific training tasks considered included flight control under turbulent air conditions. Learning, performance, and participant subjective assessment of three interface configurations were considered.

The experiment revealed differences in the inherent capabilities of three different interface configurations (a Cabin with motion turned ON, a Cabin with motion turned OFF, and a Desktop PC-based simulator) to support experiential learning given a highly complex task of Combat Search and Rescue. The complexity of the task involved maintaining aircraft control throughout a mission measured by a number of attributes. These included heading, velocity and altitude under both non-turbulent and turbulent atmospheric conditions.

Table 8. Learning Summary

Configuration	Number of Objective Measures that Support Learning	Number of Objective Measures that Do Not Support Learning
Cabin with Motion	7	5
Cabin with No Motion	9	3
Desktop	2	10

Table 8 above summarizes those findings. For the Cabin with Motion configuration, all learning measures are supported except speed and altitude control (in both turbulence and no turbulence environments) and crashes in non-turbulent environment. Learning these tasks may have increased had more runs been available to master the motion present throughout each run. For the Cabin with No Motion configuration, all measures are supported except speed control in no turbulence environment and altitude control (in both turbulence and no turbulence environments). For the Desktop configuration, only two measures are supported, the "Followed Corridor" GO/NO GO measure and altitude control in non-turbulence environment.

Table 9. Comparison against Null Hypothesis of No Learning for Each Training Configuration

Configuration	Chi-Square Test Result
Cabin with Motion	0.0070
Cabin with No Motion	0.0007
Desktop	0.4602

To gain an overall assessment of each configuration, Chi Square Tests were performed to compare the total number of objective measures that supported learning in each training configuration against the null hypothesis of no learning. The results are shown in Table 9 above. The null hypothesis of no learning is therefore rejected and alternative hypothesis that learning occurred in both Cabin configurations is accepted. Overall the null hypothesis of no learning can not be rejected for the Desktop model.

Table 10. Objective Measures that Support Learning: Statistical Comparison of Two Training Configurations

Configuration Comparison	Chi-Square Test Result
Cabin with Motion vs. Cabin with No Motion	0.665
Cabin with Motion vs. Desktop	0.0917
Cabin with No Motion vs. Desktop	0.0140

In Table 10, a Chi Square Test was performed to compare the number of objective measures that supported learning between the different systems. Learning in the Cabin with Motion configuration and the Cabin with No Motion configuration could not be statistically differentiated. Learning in the Cabin with Motion configuration was not statistically different than Desktop configuration. Learning in the Cabin with No Motion configuration was statistically different than the Desktop configuration.

11. Conclusions

No negative issues were identified with the underlying X-Plane software. Judgments by expert UH60 pilots on the suitability of the X-Plane helicopter flight model were also not negative though this does not negate the need for formal validation of flight models. Further the off-the-shelf and PC-based nature of the X-Plane software meet our upfront low cost and immediate availability interests more than the other simulation and gaming alternatives available to us at the time. Learning did occur in all training media types but the level of learning achieved in the time frame provided for learning may not be acceptable to many possible users. This is particularly true for the desktop media setup with the monitor size that we used. Many of the issues associated with the evaluation of this game for this serious training require additional future research and overlap traditional simulation research topics.

11.1 Future Research for PC, Game-based Aviation Training Systems

The requirement for a motion platform and/or a large monitor needs further research given their impact on deployment and cost. Previous research supports the argument that the learning that occurred in both Cabin configurations may have been due largely to the 60" (diagonal) rear-projection visual display system both Cabins had. The Desktop configuration used a standard 19" diagonal monitor. In their study, Reeves and Naas [34] concluded that images on a large screen (90" versus 22" diagonal) are remembered more than those in a smaller screen, provide more excitement, and more positive evaluations of the content display. Yet Reeves and Nass warned that viewers may be over stimulated by large images to the point where they may not attend to the instructional message. Tan [35] used two monitors of different size, with the same field of view. He concluded that "physical display size seems to immerse users more within virtual environments and bias users into egocentric strategies." Furthermore, he concluded that "egocentric strategies only aid performance on tasks which benefit from having users imagine their bodies within the problem space." Thus it appears that the lack of a large screen may be the dominate factor in the performance differences between the Desktop and Cabin configurations.

While this research shows learning using either Cabin treatment, the motion platform remains in question. While the motion platform did not show a statistically significant difference in overall performance from the no motion platform, a number of respondents self assessments in the no motion configuration may indicate a need for motion. This may indicate a lack of confidence in the effectiveness of the no motion platform to train a pilot for actual conditions under which motion and turbulence will be experienced. A lack of confidence may contribute to lack of use. Hence a conflict may exist between pilot performance and pilot self-assessment. Further, the quality of motion in terms of latency may also need further evaluation. 100ms latency may contribute, though not at a statistically significant level, to the dissatisfaction with the simulated pedals and aircraft movement latency expressed within the Cabin with motion participants. Discerning a more acceptable level for simulation latency of the motion platform in the presence of moderate air turbulence/microburst may require further study with multiple motion platforms with different latencies and motion delivery mechanisms.

The widespread issues with the fidelity of the terrain and feature models need further discussion. Both the Cabin with Motion and the Desktop configurations may have been adversely affected by terrain fidelity. Specifically, participants from the three training configurations commented that the terrain needed additional fidelity in order to be able to judge speed and distance. This infers that many civilian databases may need to be locally enhanced to meet mission rehearsal requirements. The lack of the ability to clearly discern this distance would adversely affect depth perception and hence the judgment of distances and speed control. Future research should focus on the appropriate fidelity of the terrain models used by the gaming community particularly when it comes to navigational landmarks for low level flight by helicopters under conditions of air turbulence and micro bursts. Aspects of the research should concentrate on determining the relationship between depth perception of the models for various altitude and speed control parameters.

As way of epilogue, Bristow Academy has since adopted a Bell 206 flight simulator as part of their training program and none of the three generic-game approaches considered above. By choosing a focused-simulator over the less-expensive game approaches, Bristow Academy's actions undermines the general notion that generic, commercial game technology is currently disrupting existing flight simulation paradigms. Whether the current generic nature of a game and game interfaces, which makes them widely popular and low-cost, is a fundamental flaw for games as a disruptive technology has yet to be determined. At the same time, specialization in games and increased fidelity of game interfaces are becoming more widespread while controls in actual fielded aviation systems are often taking on game interface usability attributes. Hence the apparent direction of simulation is toward games and games toward simulation.

12. References

[1] Abt, C. (1970) *Serious Games.* New York. Viking Press.

[2] Smith, Roger (2006). Technology Disruption in the Simulation Industry. *Journal of Defense Modeling and Simulation.* Volume 3, Issue 1, pages 3 – 10.

[3] Christensen, C. *The innovator's dilemma: when new technologies cause great firms to fail.* Boston: Harvard Business School Press; 1997.

[4] Federal Aviation Administration (1991). *Advisory Circular 120-40B: Airplane Simulator Qualification*. U.S. Department of Transportation, Washington, D.C.

[5] Federal Aviation Administration (1994). *Advisory Circular 120-63: Helicopter Simulator Qualification*. U.S. Department of Transportation, Washington, D.C.

[6] Department of Defense (2003) *Directive Number 5000.1 The Defense Acquisition System*, May 12, Washington D.C.

[7] Meyer, Austin. *X-Plane.* Laminar Research, available at http://www.x-plane.com/about.html

[8] Stewart, J.E. III, Dohme, J.A., & Nullmeyer, R.T. (2002) U.S. Army Initial Entry Rotary-Wing Transfer of Training Research. *The International Journal of Aviation Psychology*, 12(4), 359-375.

[9] Kelly, Lloyd L. (1970), *The Pilot Maker*, Grosset & Dunlap, NY.

[10] McDowell, P., Darkin, R., Sullivan, J and Johnson, E. (2006). Delta3D: A Complete Open Source Game and Simulation Engine for Building Military Training Systems. *Journal of Defense Modeling and Simulation.* Volume 3, Issue 3, pages 143–154.

[11] Kreider, L. (2002). *Memorandum: February 11, 2002.* U.S. Department of Transportation, Federal Aviation Administration, Flight Standards District Office, New Cumberland, Pennsylvania.

[12] Roessingh, Jan J. M. (2005) Transfer of Manual Flying Skills From PC-Based Simulation to Actual Flight-Comparison of In-Flight Measured Data and Instructor Ratings. *The International Journal of Aviation Psychology,* 15(1), 67-90.

[13] Talleur, D. A., Taylor, H.L., Emanuel,, T.W., Rantanen, E. And Bradshaw, G.L. (2003). Personal Computer Aviation Training Devices: Their Effectiveness for Maintaining Instrument Currency. *The International Journal of Aviation Psychology,* 13(4), 387-399.

[14] Advisory Group for Aerospace Research and Development. (1980). *Fidelity of simulation for pilot training. (AGARD-AR-159).* France: Neuilly sur Seine: North Atlantic Treaty Organization.

[15] Caro, P. W., (1977). *Some Factors Influencing Air Force Simulator Training Effectiveness. (HumRRO-TR-77-2).* Alexandria, VA: Human Resources Research Organization.

[16] Proctor, M., Panko, M. and Donovan, S. (2004) Considerations for Training Team Situation Awareness and Task Performance Through PC-Gamer Simulated Multiship Helicopter Operations. *The International Journal of Aviation Psychology,* 14(2), 191-205.

[17] Bürki-Cohen, J., Soja, N. N., Longridge, T. (1998). Simulator Platform Motion-The Need Revisited. *International Journal of Aviation Psychology*, (3), 293-317.

[18] McDaniel, W.C., Scott, P.G., Browning, R.F. (1983). Contribution of Platform Motion Simulation in SH-3 Helicopter Pilot Training. (*Technical Report 153*) Orlando, FL: Training Analysis and Evaluation Group, Department of the Navy.

[19] Hays, R.T., Jacobs, J.W., Prince, C., & Salas, E. (1992). Flight Simulator Training Effectiveness: A Meta-Analysis. *Military Psychology*, 4(2), 63-74.

[20] Rathmell, R.A. (1999). A Coalition Force Scenario 'Binni - Gateway to the Golden Bowl of Africa'. *Proceedings of the International Workshop on Knowledge-Based Planning for Coalition Forces* (ed. Tate, A.), 115-125. Edinburgh, Scotland.

[21] Allsopp, D.N., Beautement, P., Bradshaw, J.M., Durfee, E.H., Kirton, M., Knoblock, C.A., Suri, N., Tate A., & Thompson, C.W. (2002, May/June). Coalition Agents Experiment: Multi-Agent Co-operation in an International Coalition Setting. Special Issue, Knowledge Systems for Coalition Operations (KSCO), *IEEE Intelligent Systems*, 17(3), 26-35.

[22] Department of the Army (2000b) *Aircrew Training Manual Utility Helicopter, UH-60/EH-60. (TC 1-237)*. Washington, DC.

[23] Joint Chiefs of Staff. (1998, March). *Doctrine for Joint Combat Search and Rescue (Joint Publication 3-50.2).*

[24] Joint Chiefs of Staff (2002, July) *Universal Joint Task List (CJSCM 3500.04C).*

[25] Department of the Army (2000a) *Mission Training Plan for the Utility Helicopter Battalion (ARTEP 1-113-MTP)*. Washington, DC.

[26] Department of the Army (2000c). *Aeromedical Training for Flight Personnel. (FM 3-04.301).* Washington, DC.

[27] Department of the Army (2001). *Mission Training Plan for Aviation Brigades (ARTEP 1-111-MTP)*. Washington, DC.

[28] Department of the Army (2002). *Soldier's Manual and Trainer's Guide MOS 93P Aviation Operations Specialist Skill Level (STP 1-93P1-SM-TG).* Washington, DC.

[29] Department of the Army (2003). *Aviation Brigades (FM 3-04.111*). Retrieved March 25, 2004 from http://www.adtdl.army.mil

[30] Department of the Army (2005). *Aircrew Training Manual Utility Helicopter H-60 Serie (TC 1-237)*. Washington, D.C.

[31] USA Combined Arms Center. (2002, December). *Army Universal Task List (FM 7-15*). Ft. Leavenworth, KS.

[32] US Army Training and Doctrine Command (n.d.). *Command Aviation Company (CATS 01108A000*). Retrieved March 25, 2004 from http://www.adtdl.army.mil

[33] Humanalysis, Inc. (1994). *Cues and Conditions for UH-60 Blackhawk Helicopter Flight and Mission Tasks Performed by Pilots and Co-Pilots*. Orlando, Fl: US Army Simulation, Training and Instrumentation Command.

[34] Reeves, B. and Nass, C. (1998). *The media equation: How people treat computers, television, and new media like real people and places.* New York: Cambridge University Press.

[35] Tan, D. (2004). *Exploiting the Cognitive and Social Benefits of Physically Large Displays. (CMU-CS-04-154*). Carnegie Mellon University.

Acknowledgements

The authors would like to thank Dr. Neal Finkelstein and staff of the US Army Research, Development and Engineering Command, Simulation and Training Technology Center and Bristow Academy, formerly Helicopter Adventures Inc of Titusville, Florida, for their sponsorship and support of this research.

Author Biographies

Michael Proctor, LTC (Retired), Ph.D. IE, CMSP, currently is an Associate Professor with the University of Central Florida Industrial Engineering and Management Systems Department as well as with the UCF Interdisciplinary Modeling and Simulation program. His research interests include games for training, Interactive Simulation, Real-Time Simulation Agents, and Simulation-Based Life-Cycle Engineering.

Maria Bauer is a Science and Technology Manager for the US Army RDECOM Simulation and Training Technology Center. She holds a Ph.D. and a M.S. in Industrial Engineering from the University of Central Florida and a B.S. in Electrical Engineering from the University of Miami. Her professional experience includes fifteen years of software engineering and five years of program management working with Department of Defense (DOD) military acquisition systems and simulation technology in Army programs. Her research interests include flight simulation, virtual reality, constructive simulations, and high performance computing.

Thomas Lucario is a former Army Aviation Officer (Major) and is currently the head of operations for technology and intellectual property holding company based in Austin Texas. He holds a BS from the United States Military Academy at West Point and a Masters in Industrial Engineering (Simulation and Training) from the University of Central Florida. His professional experience includes over eleven years as an Army Aviation officer, pilot, leader and trainer; including three years as an Army Simulations Operations Officer with time served as a Combat Aviation Brigade Simulations Officer in Iraq.