

I

**AN EVALUATION OF TRAINING INTERVENTIONS AND COMPUTED
SCORING TECHNIQUES FOR GRADING A LEVEL TURN TASK
AND A STRAIGHT IN LANDING APPROACH ON A
PC-BASED FLIGHT SIMULATOR**

**By
Bruce E. Heath**

**A Thesis Submitted to the Graduate Faculty
of Tuskegee University
in Partial Fulfillment of the Requirements
of the Degree:**

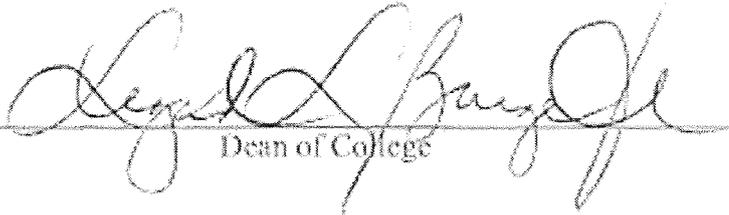
MASTER OF SCIENCE IN MECHANICAL ENGINEERING

**TUSKEGEE UNIVERSITY
Tuskegee, Alabama 36088
May 2007**

THESIS APPROVED BY:



Major Professor



Dean of College



Dean of Graduate Programs

FOR:

Bruce E. Heath

Student

**AN EVALUATION OF TRAINING INTERVENTIONS AND COMPUTED
SCORING TECHNIQUES FOR GRADING A LEVEL TURN TASK
AND A STRAIGHT IN LANDING APPROACH ON A
PC-BASED FLIGHT SIMULATOR**

Title of Thesis

**AN EVALUATION OF TRAINING INTERVENTIONS AND COMPUTED
SCORING TECHNIQUES FOR GRADING A LEVEL TURN TASK
AND A STRAIGHT IN LANDING APPROACH ON A
PC-BASED FLIGHT SIMULATOR**

**By
Bruce E. Heath**

**A Thesis Submitted to the Graduate Faculty
of Tuskegee University
in Partial Fulfillment of the Requirements
of the Degree:**

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

**TUSKEGEE UNIVERSITY
Tuskegee, Alabama 36088
May 2007**

TABLE OF CONTENTS

LIST OF FIGURES	5
LIST OF TABLES	7
ACKNOWLEDGMENTS	8
ABSTRACT	9
 Chapter	
I. INTRODUCTION	11
Motivation	11
Objective	12
Literature Review	13
Above Real-Time Training (ARTT)	13
Tunnel in the Sky	14
Performance Feedback	18
II. LEVEL TURN EXPERIMENT	20
Setup	20
Hardware	20
Software	22
Method	23
Participants	23
Experimental Design	24
Procedure	24
Orientation to Flight Controls and Aircraft Instruments	25

	3
Familiarization Flights	26
Training Flights	26
Training Briefing Scenario	27
Feedback in Training	28
Evaluation Task	28
Testing Scenario	29
Performance Metrics	29
Familiarization Flight Performance Metric	29
Turn Task Performance Metric	30
Results and Discussion	31
III. LANDING TASK EXPERIMENT	34
Setup	34
Hardware.....	34
Software	35
Method	36
Experimental Design	36
Participants	36
Procedure	37
Orientation to Flight Controls and Aircraft Instruments	37
Familiarization Flights	40
Training Flights	41
Feedback in Training	42

	4
Evaluation Flights	42
Observation	42
Performance Metrics	43
Metric Definition	43
Starting Error	45
Results and Discussion	47
IV. COMPARISON OF COMPUTED METRIC WITH CFI EVALUATION	52
Introduction	52
Method	52
Results and Discussion	53
V. CONCLUSIONS AND SUMMARY	58
APPENDICES	61
A. Visual Basic Scoring Module	62
B. Consent Form and Participant Demographic and Data Form Turn Study	65
C. Participant Demographic and Data Form Landing Study	68
D. Instructions and Data Form for Flight Instructors	71
E. Instructor Experience	74
REFERENCES	75

LIST OF FIGURES

Figure	Page
1. An example of tunnel in the sky	15
2. Straight tunnel components	17
3. LiteFlite partial mock cockpit setup	20
4. HUD for the LiteFlite software	21
5. LiteFlite Head Down Display (HDD)	22
6. LiteFlite HUD and HDD	25
7. Deviations of altitude and speed	31
8. Average errors during training and evaluation flights	32
9. Partial mock cockpit	35
10. Instrument panel from Microsoft Flight Simulator 2002	39
11. Runway alignment	44
12. Runway alignment error geometry	46
13. Normalized performance of the participants in NH group	49
14. Performances of the participants in the NH group	50
15. Normalized performance of the participants in the HH group	50
16. Performance of the participants in the HH group	51
17. Average performance of the NH and HH groups	51
18. Correlation of instructor grades	53
19. Best controlled parameters as cited by the CFIs	54
20. Worst controlled parameters as cited by the CFIs	55
21. Correlations between flight instructor scores of selected flights	55

22. Correlation between composite error metric and instructor scores	56
--	----

LIST OF TABLES

Table	Page
1. Grading Criteria for Straight and Level Flight: 5000ft, 75 knots.	23
2. Grading Criteria for Straight and Level Flight: 10000 ft, 129 kts	40

ACKNOWLEDGMENTS

The author would like to thank his father and mother and extended family for the many good thoughts and prayers. The author would like to thank his major professor, Dr. M. Javed Khan, and the thesis committee: Dr. M. Javed Khan, Dr. Marcia J. Rossi, Dr. Syed Firsat Ali Dr. Chuan-Chiang Chen and Dr. Vascar G. Harris. This author is greatly appreciative of all undergraduate students who took part in and helped with the experiments. This research was sponsored by NASA Dryden Grant NAG4-226. It was also partially supported by NASA Ames Award CTS-0320305.

ABSTRACT

AN EVALUATION OF TRAINING INTERVENTIONS AND COMPUTED SCORING
TECHNIQUES ON A LEVEL TURN TASK AND A STRAIGHT IN LANDING
APPROACH ON A PC-BASED FLIGHT SIMULATOR

By
Bruce E. Heath

One result of the relatively recent advances in computing technology has been the decreasing cost of computers and increasing computational power. This has allowed high fidelity airplane simulations to be run on personal computers (PC). Thus, simulators are now used routinely by pilots to substitute real flight hours for simulated flight hours for training for an aircraft type rating thereby reducing the cost of flight training. However, FAA regulations require that such substitution training must be supervised by Certified Flight Instructors (CFI). If the CFI presence could be reduced or eliminated for certain tasks this would mean a further cost savings to the pilot. This would require that the flight simulator have a certain level of 'intelligence' in order to provide feedback on pilot performance similar to that of a CFI. The 'intelligent' flight simulator would have at least the capability to use data gathered from the flight to create a measure for the performance of the student pilot. Also, to fully utilize the advances in computational power, the simulator would be capable of interacting with the student pilot using the best possible training interventions.

This thesis reports on the two studies conducted at Tuskegee University investigating the effects of interventions on the learning of two flight maneuvers on a flight simulator and the robustness and accuracy of calculated performance indices as compared to CFI

evaluations of performance. The intent of these studies is to take a step in the direction of creating an 'intelligent' flight simulator. The first study deals with the comparisons of novice pilot performance trained at different levels of above real-time to execute a level S-turn. The second study examined the effect of out-of-the-window (OTW) visual cues in the form of hoops on the performance of novice pilots learning to fly a landing approach on the flight simulator. The reliability/robustness of the computed performance metrics was assessed by comparing them with the evaluations of the landing approach maneuver by a number of CFIs.

CHAPTER I

INTRODUCTION

Motivation

One result of the relatively recent advances in computing technology has been the decreasing cost of computers yet increasing computational power. This has allowed high fidelity airplane simulations to be run on personal computers (PC). Thus, simulators are now used routinely by pilots to substitute real flight hours for simulated flight hours for training for an aircraft type-rating thereby reducing the cost of flight training. However, FAA regulations¹ (quoted below) require that such substitution training must be supervised by Certified Flight Instructors (CFI).

5. AUTHORIZED USE.

a. Instruction by an Authorized Instructor. Qualified PCATD's may be highly beneficial when used under the guidance of an authorized instructor to achieve learning in certain procedural tasks such as area departures and arrivals, navigational aid tracking, holding pattern entries, instrument approaches, and missed approach procedures. Accordingly, the FAA has determined to continue the policy that any time instruction is to be used to log time toward meeting any requirement of the regulations, an authorized instructor must have presented the instruction.

If the CFI presence could be reduced or eliminated for certain tasks, this would mean a further cost savings to the pilot. This would require that the flight simulator have a certain level of 'intelligence' in order to provide expert feedback on pilot performance similar to that of a CFI. The 'intelligent flight simulator' (IFS) would have at least the capability to use data gathered from the flight to create a measure for the performance of the student pilot and provide expert feedback to the student pilot based on the computed performance index. Also,

to fully utilize the advances in computational power, the simulator would be capable of interacting with the student pilot using the best possible training interventions.

Ludwig². et al. have referred to Mulgand's³ main components of an Intelligent Tutoring System (ITS) to be the 'helper' and the 'advisor'. In their implementation of an IFS, the 'helper' assisted the pilot by making the helicopter, at first, easier to fly. As the student became more comfortable with the aircraft, the pilot received less help. The 'advisor' communicated directly with the pilot using text-to-speech software in four roles namely, tutorial, performance monitoring, monitoring flight control manipulation and advisory, verbalizing suggestions to control or correct flight.

Objective

The objective, however, of this study has been to investigate two important aspects of an IFS. These aspects being: (a) novel techniques to determine their efficacy in expediting the training of novice pilots and (b) reliability and robustness of computed performance indices in comparison to the CFI grading.

The two novel training techniques that were evaluated during the course of this investigation were:

- (a) Above real time training (ARTT) for training a level S-turn with and without performance feedback.
- (b) Out-of-window (OTW) visual cues based on "tunnel in the sky" for training a landing approach. Performance feedback was used.

Literature Review

Above Real Time Training (ARTT)

In 'above real time' (ART), or as it has been referred to in literature⁴ as 'time compression', events are presented in a simulator as if they were happening in real time but are actually running faster than the same event in the real world. For example, an aircraft flying at 90 knots might look as though it is flying at 180 knots. However, any airplane flying at twice the speed would have, for a level-turn, a 4 times larger radius. In fact, lift and drag for the airplane at 180 knots would be different from an airplane flying at 90 knots. ART, therefore, means that in a simulator, the aircraft flying at 90 knots will have lift, drag and airspeed indications of an aircraft flying at 90 knots, except that time will pass at a multiple of actual clock time. Thus, for example, at 2.0 ART, the airplane flying at 90 knots will cover 90 nautical miles in one hour of computer clock time, but in real clock time it will take 30 minutes. Previous experiments^{4, 5, 6} on training of pilots have indicated that real time training reinforced with ARTT could offer an effective training strategy for tasks which require significant effort at time and workload management. Kolf⁵ documented that no matter how much pre-flight simulator training the pilot had undergone, the actual flight appeared, to the pilot, to take place in a much faster time frame. He chose pilots with experience flying the M2-F3 to fly experimental flights in the simulator and noted that by increasing simulator clock speed to 1.5 ART, the pilots felt that the simulator more closely resembled their actual flying experience. Kolf hypothesized that for ARTT an appropriate above real time factor would have to be chosen and would be a function of aircraft type, individual, task, and experience. Hoey⁴ compared biomedical measurement data of test

pilots flying remotely piloted vehicles, with data taken in past flights. He suggested that providing ARTT on a simulator could approximately simulate the mental state of pilots as if actually flying. His suggestion was that ARTT compared with real time training on a simulator, would lead to a smoother transition from simulator to airplane. Crane and Guckenberger⁶ reference Schneider's⁷ air traffic control study who proposed that "the primary effect of time compression is to allow more training trials within a given period of clock-time." Crane and Guckenberger⁵ contrasted this study by giving two groups the same number of trials so that the above real-time participants receive less clock-time than the real-time participants while performing the task of tracking and shooting at a target in the air. The test trials were then performed in real-time. The results of that study showed that participants trained in above real-time performed better than participants trained in real-time.

However, this review of literature on ARTT determined that ARTT has not been exploited for the training of flight maneuvers.

Tunnel in the sky

The second training intervention investigated was the use of visual cues. This method is based on the concept of Tunnel in the Sky. The tunnel in the sky (Figure 1) which is a visual aid that shows the proposed path of flight for a maneuver in the sky is the subject of a number of studies.^{8,9} The benefits of using tunnel in the sky for navigation versus flight path tracking with conventional instruments have been investigated in these studies.

Barrows, Alter, Enge, Parkinson, and Powell¹⁰ compared the use of a tunnel display versus conventional instruments. In the simulator three different displays were offered:

1. The control group: conventional instruments, vertical speed indicator (non-instantaneous), horizontal situation indicator (HSI) with glide slope display, attitude indicator, altimeter, airspeed indicator, and turn coordinator,
2. First experimental group: Conventional instruments with a track symbol added to the HSI.

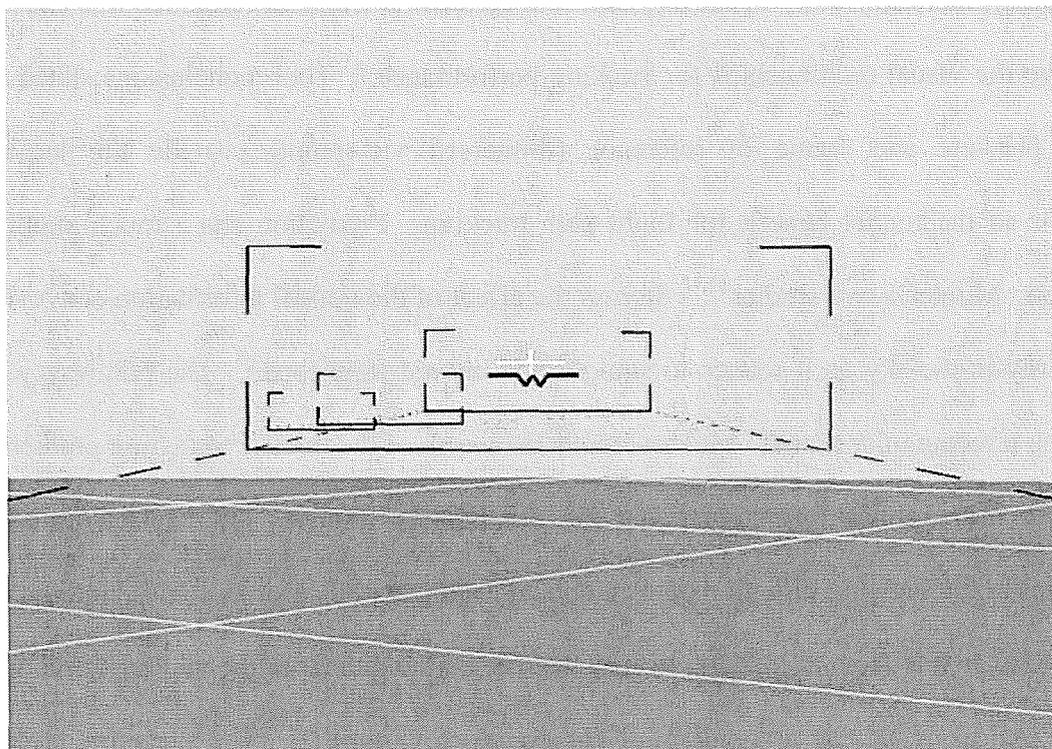


Figure 1. An example of tunnel in the sky.

3. Second experimental group: The HSI from the first experimental group and an additional symbol “glide slope predictor”.
4. Third experimental display: a pathway-in-the-sky.

Using a simulator and actual flight, they found that the tunnel display offered significant improvements in horizontal and vertical flight precision and in workload reduction as compared to conventional instrumentation.⁷

Doherty and Wickens⁹ looked at the effects of preview, prediction, frame of reference and display gain in tunnel in the sky displays. In short, preview allows a pilot to anticipate upcoming demands and prepare a response before actually starting a maneuver. They also used a flight predictor in their display that reduced the cognitive effort required by the pilot in determining future trajectory of the airplane. An egocentric frame of reference was chosen to “preserve the visual relationships for tracking performance.” Their conclusions stated in part that preview and frame of reference (immersed viewpoint) provide the largest contribution to the tunnel benefit for flight path tracking. Prediction has a much smaller contribution. Mulder’s two studies^{11,12} discuss the effect of the optical information conveyed by a pathway-in-the-sky as it relates to straight and curved trajectories. The pathway was drawn using a rectangular tunnel. Mulder credits the worth of the tunnels to texture gradients that provide the useful information about the participants’ motion. In Mulder’s study, parallel lines to the viewing direction convey optical “splay angle” information or gradient of perspective (lines that meet at infinity). Lines that are perpendicular to the viewing direction convey an optical density or a “gradient of compression” (Figure 2). The study found that for a straight tunnel, optical splay and optical density were “essential for the pilot to perceive and control the aircraft position and motion with respect to the tunnel.” Also, Mulder found that splay and density were complementary. A tunnel display without optical splay lines is not recommended and a tunnel display with the wrong number of frames could cause clutter.

For a curved trajectory, gradients have a much smaller use. First of all there is no vanishing point as can be seen in a straight tunnel. Second, only the gradients conveyed by the nearby elements of the tunnel provide valuable information about the aircraft movement.

Third, the presented curved trajectory does not convey actual position and attitude relative to the trajectory. These are presentation biases that lead the pilot to compensate for position and attitude errors that are not really there.

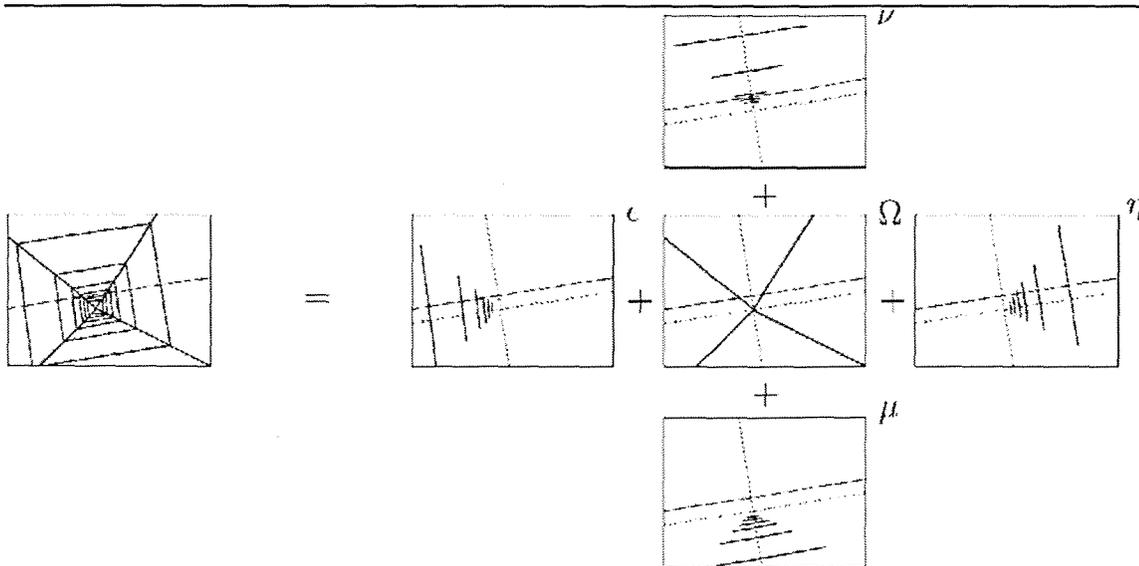


Figure 2. Straight tunnel components. Ω shows the optical splay angle, ϵ, ν, η, μ show the perpendicular lines that convey optical density (Mulder.¹¹).

For a curved trajectory it was found that splay lines are substantially but not significantly useful. Presentation bias is greatest when the display shows only contour lines and decreases considerably when tunnel frames are used. Presentation bias leads to errors in the lateral position relative to the center circle. Participants found it difficult to tell when they were flying on the center of the path. Also, Mulder recommends use of a greater number of rings to reduce presentation bias.

Performance Feedback

As mentioned above, the cost component associated with the requirement of a CFI to be present during every simulation can be minimized if the flight simulator is able to emulate expert feedback of a CFI. Vruels and Obermayer¹³ emphasize the need of automated performance measures on modern simulators as a substitute for direct evaluation of performance by an instructor. Rantanen and Talleur¹⁴ have suggested the following five primary measures for pilot performance from the flight data recorder for nine flight parameters: (1) Standard deviation. A small standard deviation is indicative of good performance. (2) Root mean square error (RMSE) which summarizes the overall error. (3) Number of deviations, which is a measure of occurrences of the aircraft staying outside the prescribed tolerances. (4) Time outside tolerance is the cumulative time the aircraft spends outside prescribed tolerances. (5) Mean time to exceed tolerance is computed from the rate of change between successive data points and the aircraft's position relative to a given tolerance. Heath and Crier¹¹ used computerized scoring and certified flight instructors (CFI) to attempt to devise a computed performance measure for a level turn. A performance index based on deviations from the required heading, altitude and airspeed was compared with the instructors' letter grades. One of the challenges of such comparisons is inter-rater reliability of the flight instructors. In a study about crew resource management (CRM) in a simulated helicopter, Brannick, Prince, and Salas¹⁵ evaluated three items: "(a) specific crew behaviors in response to scenario events (e.g., whether the crew kept out of icing conditions), (b) valuations of crew responses to scenario events (e.g., overall handling of the icing problem),

and (c) crew resource management (CRM) dimensions for the entire scenario (e.g., evaluations of decision making).”

They were able to show that “respectable levels of both interjudge agreement and internal consistency were achieved for items dealing with the evaluation of crew performance in response to items in events embedded in the scenario”. Interjudge agreement was low but internal consistency was high on CRM items and scales. Interjudge agreement was high but internal consistency low for specific observable behaviors.

An attempt to correlate instructors’ grades with a performance metric by Heath and Crier¹⁶ showed the same result. The instructors did not agree on how to measure performance. Discrepancies between instructors may be due partially to the fact that instructors normally give a pass/fail type of evaluation and not the type of A, B, C, D, E grade requested for the current study.

In a follow-up study the instructors were asked to identify the parameter they thought the student best controlled and the parameter they thought was worst controlled. Performance metrics were graphed against the grades given by the three instructors for the flights where they were in agreement of the best and worst controlled parameters. Using that kind of selection, a better correlation was found between instructor grades.

CHAPTER II

LEVEL TURN EXPERIMENT

This chapter describes a study which compared Above Real Time Training (ARTT) versus Real Time Training to investigate the performance of 32 novice trainees using a Level Turn Maneuver. The turn flight maneuver is one of the basic flight maneuvers required for a pilot in pre-solo training. According to the Federal Aviation Regulations other basic flight maneuvers include takeoff and landing, straight and level flight, climb and descent. (FAR 2002; §61.87) The purpose of this experiment is to investigate the effect of different levels of ARTT on novice pilots and to attempt to create a performance metric that will allow the comparison of the various groups of students flying at different levels of ART.

Setup

Hardware

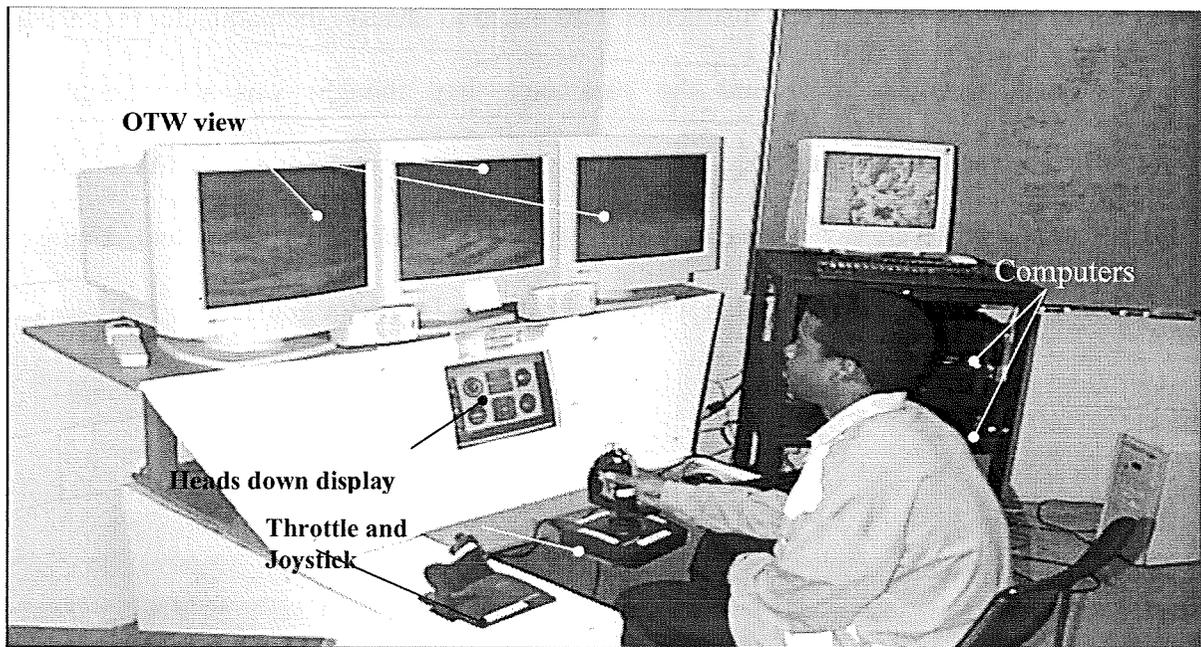


Figure 3. LiteFlite partial mock cockpit setup shows heads down display and OTW.

The experiment was performed in the Flight Vehicle Design Lab at Tuskegee University. The “simulator” consists of a mock setup of a partial construction of a cockpit. Three 19” monitors create a panoramic, outside of the window (OTW) view. The center monitor contains a heads up display (HUD) superimposed on the forward view. The HUD displays a composite of several relevant instruments including: altimeter, airspeed indicator, horizontal situation indicator (HSI), turn and slip coordinator, and heading indicator as shown (Figure 4).

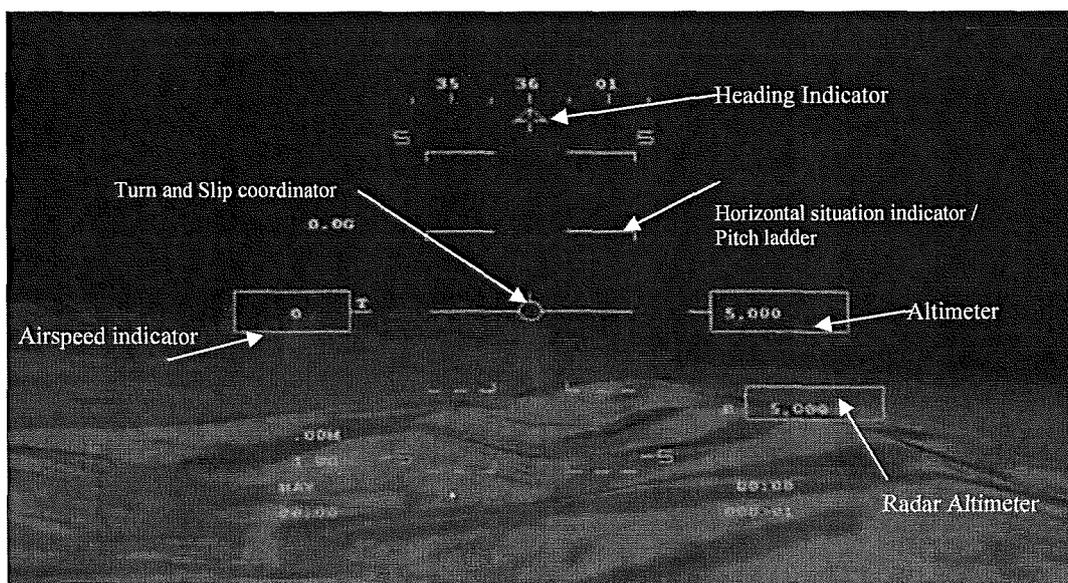


Figure 4. HUD for the LiteFlite software is superimposed on the center monitor of the OTW view.

The mock setup had an inside the cockpit “instrument panel” view displayed on a 15” monitor as shown (Figure 5).

The joystick used was a Saitek X36F and the throttle was a Saitek X35T controller. Rudder pedals used were from CH products. The PCs for the simulator were two Heavy Metal computers made by Quantum 3D.

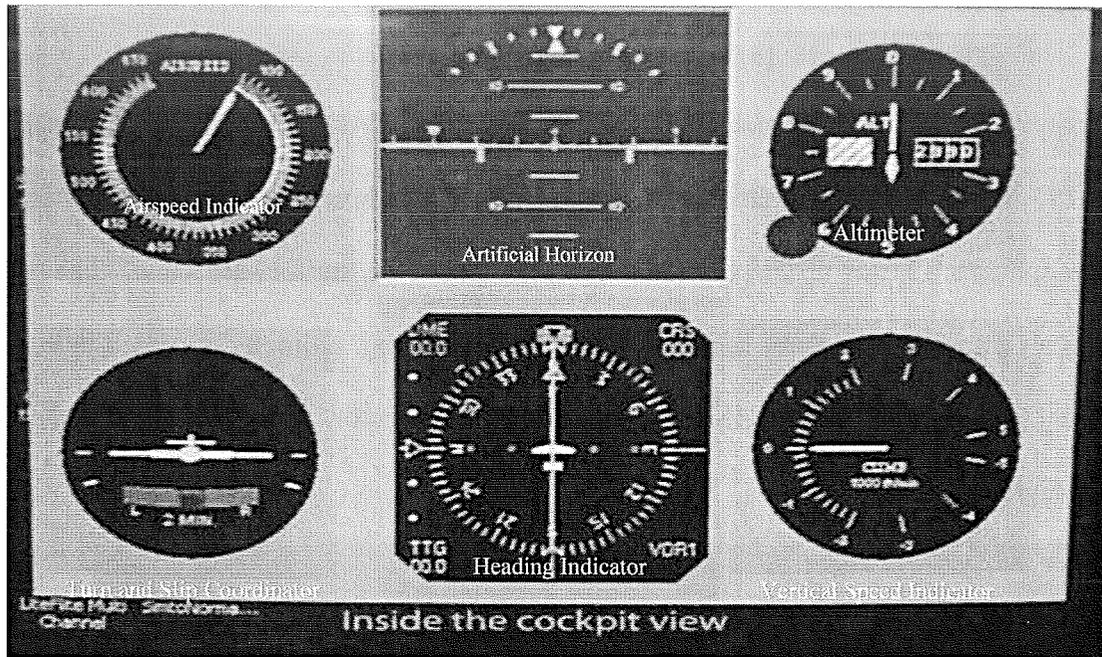


Figure 5. LiteFlite Head down display (HDD).

Each computer has 2 Pentium II processors running at 400Mhz, 400MB RAM, and has three extra video cards for the Out of the Window (OTW) view. Figure 3 also shows a moving map display monitor on the right of the picture. The moving map display was not used in this experiment. Additional details can be found in Williams¹⁷.

Software

Flights were conducted using the LiteFlite version 3.3 flight simulation software developed by SDS International, Orlando, Florida. LiteFlite offers the Predator (UAV) simulation which was used in these studies. LiteFlite has an internal facility to capture flight data using Visual Basic. The same feature exports the data to *.csv files that are readable by Microsoft's Excel spreadsheet program.

Method

Participants

The student participants for this study were 36 undergraduate freshmen at Tuskegee University who had little or no prior experience with an airplane simulator. All participants enrolled in an Introductory Psychology class. The student participants were offered extra course credit in their psychology course for their participation. Our acceptance of a participant into the experiment, beyond the familiarization phase, was based upon their adherence to the flight parameters described in Table 1. Participants who achieved a score of 2.0 (grade based on comparing parameter values in the flight with the parameter values and tolerances required) on at least one familiarization flight were allowed to continue into training. The grading scheme is explained in Table 1. Thirty two participants who achieved the passing grade continued on to the training phase. Each of the participants who achieved a score of 2.0 was randomly placed in one of six groups.

Table 1. Grading Criteria for Straight and Level Flight

Nominal Altitude = 10000 ft	Nominal Speed = 129 knots	Nominal Bank angle = 0 degrees	Grade
Deviation(Δ in altitude)	Deviation(Δ in speed)	Deviation(Δ in degrees)	
$\Delta = \pm 50$	$\Delta = \pm 5$	$\Delta = \pm 3$	A = 4.0
$+50 \leq \Delta \leq +100$	$+5 \leq \Delta \leq +10$	$+3 \leq \Delta \leq +6$	B = 3.0
$-100 \leq \Delta \leq -50$	$-10 \leq \Delta \leq -5$	$-6 \leq \Delta \leq -3$	
$+100 \leq \Delta \leq +150$	$\pm 10 \leq \Delta \leq \pm 15$	$+6 \leq \Delta \leq +9$	C = 2.0
$-150 \leq \Delta \leq -100$	$-15 \leq \Delta \leq -10$	$-9 \leq \Delta \leq -6$	
$+150 \leq \Delta \leq +200$	$\pm 15 \leq \Delta \leq \pm 20$	$+9 \leq \Delta \leq +12$	D = 1.0
$-200 \leq \Delta \leq -150$	$-20 \leq \Delta \leq -15$	$-12 \leq \Delta \leq -9$	
$+200 < \Delta$ or $\Delta < -200$	$+20 < \Delta$ or $\Delta < -20$	$+12 < \Delta$ or $\Delta < -12$	E = 0.0

Experimental Design

This experiment used a 3X2 factorial design. There were three types of training used: all flights flown in real time (RT), the first half of the flights flown in 1.5 above real time and the second half flown at 2.0 above real time (ART), or the first half of the flights flown at 2.0 ART and the second half of the flights flown at 1.5ART. Half of the participants in each group were randomly assigned to receive feedback or no-feedback during training. Feedback for this study took the form of the student instructors making suggestions to the participant as to how to improve the handling of the airplane. For the feedback group, students were given a picture representation of their performance in the form of a ground track on the second, fourth, sixth, eighth, and tenth flights. For the study conducted on the flight simulator, all participants experienced the following segments: Orientation to Simulator Controls and Functions, 2 Demonstration Flights, 3 Familiarization Flights, 8 Training Flights, and 2 Evaluation Flights. The evaluation flights were all conducted in RT.

Procedure

All participant pilots experienced the following sessions: Orientation to flight controls and aircraft instruments, familiarization flights, training flights and evaluation flights.

The instructors individually attending the participants in this experiment were the author, from the Aerospace Science Engineering Department, and three other juniors and seniors in the Psychology Department at Tuskegee University. The undergraduate instructors had, for the most part, no flying or simulator experience so they spent a good deal of time becoming comfortable with the controls and functions of the simulator. The graduate

instructor has 10 hours of flight time in a Cessna 172 and numerous hours using Microsoft Flight Simulator.

Orientation to Flight Controls and Aircraft Instruments

A protocol was formulated to introduce the novice subjects to the flight controls. Such a protocol was considered essential so as to ensure standardized briefing to all participants thereby minimizing the influence of different instructors. During the orientation process, participants were shown a model airplane and informed of the basic control surfaces of the aircraft, their functions, and the movements associated with each control surface. The participants were then given an overview of the locations and functions of the joystick, rudder pedals, and the throttle located in the mock cockpit. Then, participants were instructed in the location and functions of the following displays on the HUD (head up display): altimeter, radio altimeter, airspeed indicator, heading tape, artificial horizon, pitch ladder, and clock.

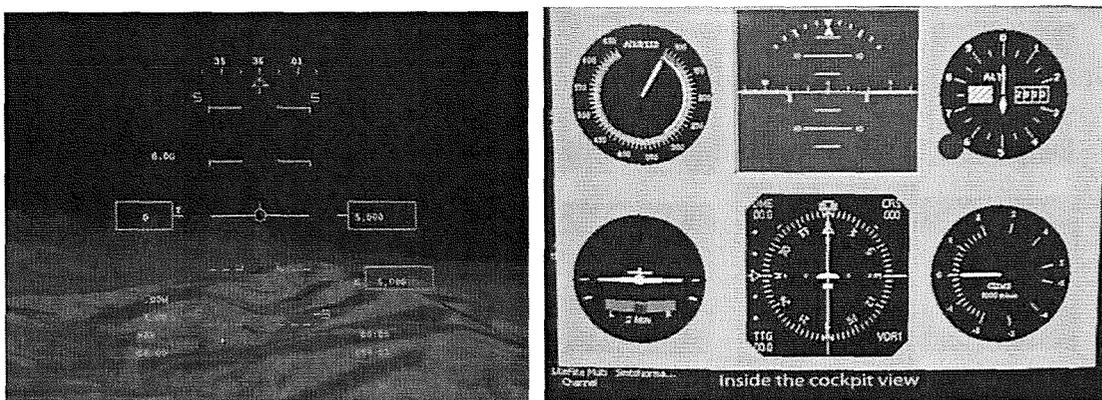


Figure 6. LiteFLite HUD and HDD.

Participants were also instructed on the location and functions of the following instruments on the HDD (head down display): artificial horizon and the vertical velocity

indicator. Because the turn-and-slip indicator on the HDD was not functioning properly, participants were explained the use of the HUD to determine if the aircraft was making a coordinated turn.

Familiarization Flights

Participants observed the experimenter flying one 3-minute straight and level flight while explaining the controls and operation of the aircraft. Participants flew four flights of 3-minutes duration each in which they were told to maintain airspeed of 129 knots, a heading of 360 degrees, and an altitude of 10,000 feet. The participants were coached on all but the last of the flights. On the first flight, the experimenter asked the student to focus on altitude maintenance primarily. On the second flight the experimenter asked the student to focus on altitude and heading. For the third flight the experimenter asked the student to maintain altitude, heading and speed. The experimenter assisted with the controls, if needed. At the end of each of the flights participants were shown their actual course and given feedback on how well they performed. On the last flight, the participant was to maintain altitude, heading and airspeed without being coached.

During the familiarization flights, participants were graded on a 4.0 grade scale. The grades were calculated by a program written in visual basic by SDS.

Training Flights

The experiment consisted of a training phase and an evaluation phase. In the training phase pilots flew a mission consisting of a coordinated 180° turn with a 10° bank angle while the evaluation task was an S-turn with a bank angle of 30° . Each group conducted their

training in the following sequence immediately followed by two evaluation flights in real time. Participants in Groups II, IV and VI were given post-flight feedback (PFFB).

Group I: Five missions in 1.5 ARTT then five missions in 2.0 ARTT with No-PFFB,

Group II: Five missions in 1.5 ARTT then five missions in 2.0 ARTT with PFFB,

Group III: Five missions in RTT then five missions in RTT with No-PFFB,

Group IV: Five missions in RTT then five missions in RTT with PFFB Feedback,

Group V: Five missions in 2.0 ARTT then five missions in 1.5 ARTT with No-PFFB,

Group VI: Five missions in 2.0 ARTT then five missions in 1.5 ARTT with PFFB.

Training Flights Briefing Scenario

To provide some realism to the training, the participants were briefed on their task according to the following scenario:

“You’re the flight leader of a combat air patrol on a routine mission at 5000 ft altitude, 129 knots and heading in a 360° direction. An unidentified low-speed aircraft is being tracked by radar moving towards a no fly zone. Despite repeated requests, the aircraft fails to identify itself and continues towards the no fly zone. The Air Defense Radar Controller assigns the interception to your flight in order to make a visual identification. You are asked to initiate a right hand turn, maintain a bank angle of 10 degrees, an altitude of 5000 ft and a speed of 129 knots. Exactly after two minutes (as recorded on your heads up display), you should level out. Your heading should be 180°, putting you right behind the target, close enough for visual contact. You will complete a number of these missions. Do your best, because once you have finished the practice missions, you will be required to complete a test mission with a slightly more difficult task.”

Feedback in Training

Participants in the No-PFFB groups received no feedback during or after flights. If they asked questions about the task, they were re-read the relevant instructions. Participants in the PFFB groups received verbal feedback after each flight. In verbal feedback, the experimenter told the participant the errors in manipulating the controls and in focusing on the wrong instruments. After the second, fourth, sixth, eighth, and tenth flights, participants in the PFFB groups were also shown a screen printout of the desired ground track and the pilot's actual ground track for comparison. Then, the experimenter explained possible reasons for the deviations in performance.

Evaluation Task

The evaluation flight was an S-turn. Pilots were instructed to make a 30° banked right turn to 180° or south and then an immediate 30° banked turn to 0° or north. The participants were told that they would fly two more flights having a different and more challenging task. Both flights were conducted in real-time for all participants and lasted for three minutes each. No feedback or coaching was provided during testing. Participants were read the following scenario before each flight. The participants were to fly an S-Turn with 30° bank angles, maintain 5,000 ft. and a speed of 164 knots.

However, a loss of situational awareness was observed in most of the participants after the first leg of the S-turn during evaluation. Those people who lost situational awareness could be seen doing some of the following: continuing the turn in the same direction, banking the plane and then pulling the stick back to increase the rate of turn, or

continuing in a straight line after the first turn. Thus, only the first leg of the S-turn was evaluated for analysis purposes.

Testing Scenario

The participants were given the following scenario for evaluation flights:

“As part of a Combat Air Patrol, the Air Defense Radar Controller notifies you that a hostile aircraft is being tracked heading towards a vulnerable point (VP). Your instructions are to initiate a right hand turn in order to pursue the aircraft while maintaining a bank angle of 30° , an altitude of 5000 ft., and a speed of 164 knots. Exactly 48 seconds after the instructions to initiate the turn you should level out. You will now be exactly behind the target on a heading of 180° . You are then informed of a second target. You should fire your missile and immediately initiate a left turn with a 30° bank maintaining 5,000 ft. altitude and a speed of 164 knots in pursuit of the second target assigned. Exactly 48 seconds after initiating the second turn you should level out. Your heading should be 360° , putting you right behind the second target. Launch your missile. Mission accomplished.” (It should be noted that no missile function was provided in the simulation).

Performance Metrics

Familiarization flight performance metric

The performance metric for the level turn was a letter-grade scheme based essentially on grading scheme of Williams.¹⁷ The same scheme was used for selecting participants for the study (Table 1).

Turn Task Performance Metrics

The performance metric to track learning was calculated. This performance measure was based upon an aggregate of non-dimensional altitude error and the error, in radians, for the bank angle. The two measures were then squared, and the square root was taken of that result. Hence:

$$\text{Error} = \sqrt{(\Delta H / \mathfrak{R})^2 + \Delta\phi^2}$$

where the radius of turn 'Y' was calculated for the flight parameters to be maintained by the pilot using the following relation¹⁶:

$$Y = \frac{V^2}{g \cdot \sqrt{n^2 - 1}} ; n = \frac{1}{\cos\phi}, \text{ g = acceleration due to gravity for the } V = \text{airspeed,}$$

and: ϕ = turn bank angle

Here, altitude error (ΔH) is non-dimensionalized by the theoretical radius (Y) of the turn made by the simulated airplane. If H is not divided by Y small changes in altitude overwhelm any contributions by possible large changes in bank angle (ϕ). Also, error in speed was not used in this calculation because the pilots did not make use of the throttle. Hence, as altitude decreased, speed increased and vice versa. The effect of speed changes was thus captured in the error in altitude as can be clearly seen in Figure 7.

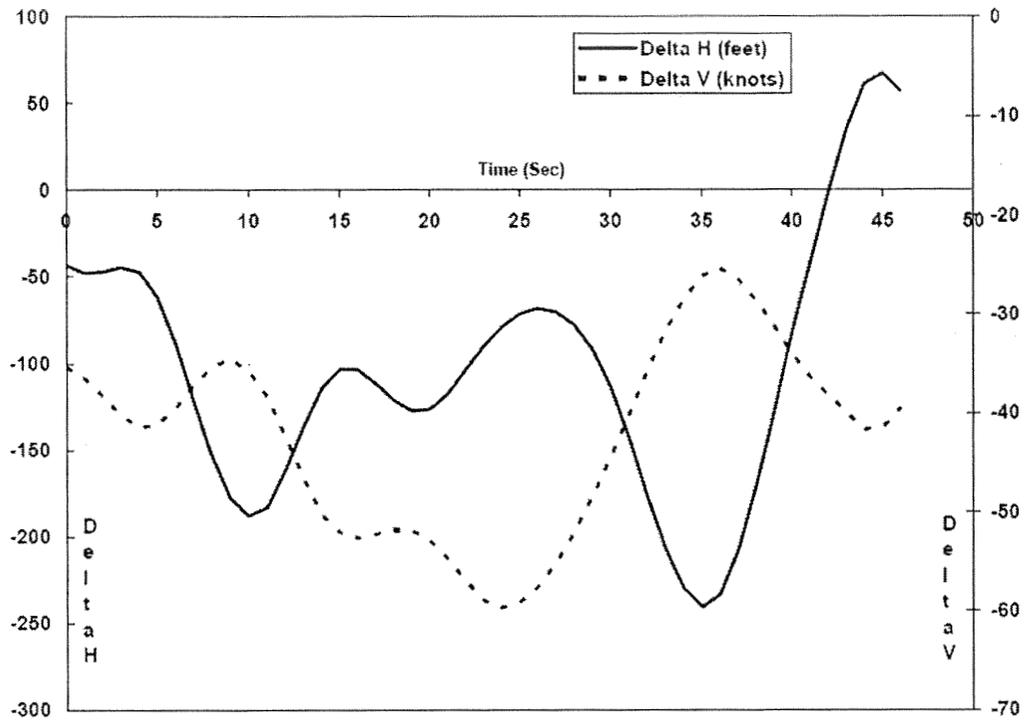


Figure 7. Deviations of altitude and speed (typical for a single flight).

Results And Discussion

The graph in Figure 8 shows the average errors for each group during training and evaluation flights. For the ART training regimen the first five flights and last five flights were at either 1.5 above real time or 2.0 above real time. In either case after the fifth flight the one level of above real time was changed to the other.

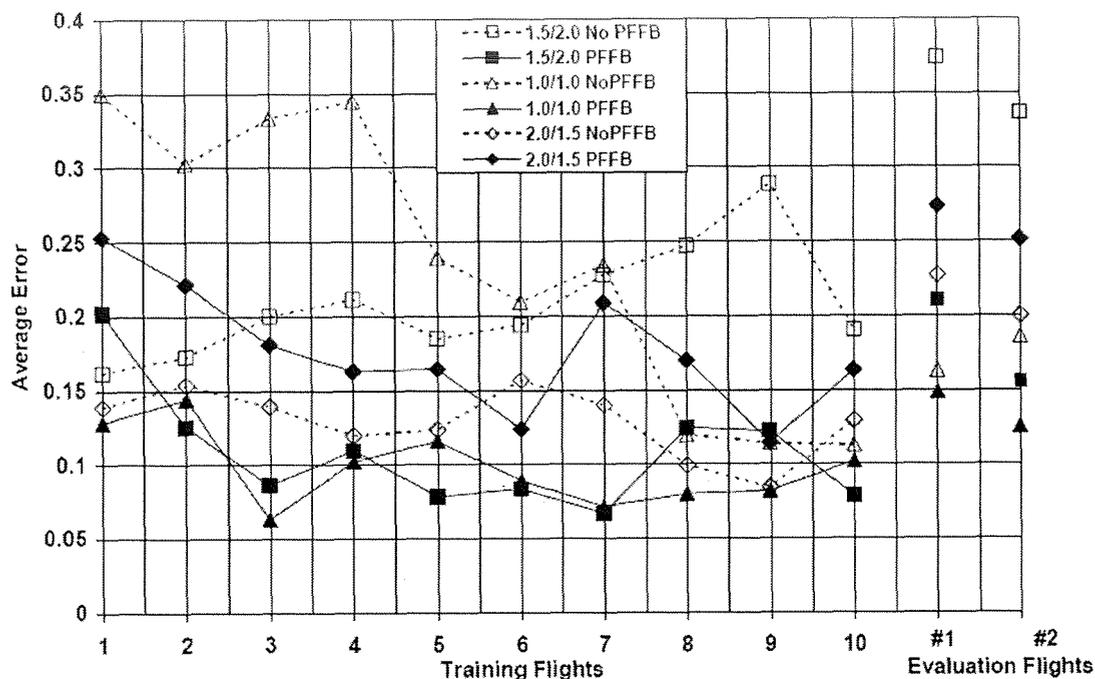


Figure 8. Average errors during training and evaluation flights.

In terms of post flight feedback or no post flight feedback, the graph shows that for both 1.5/2.0 ART and 1.0/1.0 RT, the No-PFFB group did worse than the corresponding PFFB group. This was expected because each participant in the PFFB group was given the advantage of knowing how well or poorly he was doing. For the 2.0/1.5 group the No-PFFB group actually did better than the corresponding PFFB group. This second result was unexpected in that feedback should provide extra help for the participant to understand what is going wrong. This may be due to the fact that the PFFB started the experiment with a worse performance; thus perhaps there was a difference between the two groups from the start of the experiment.

It might be expected that for the PFFB and No-PFFB groups, the performance of the 1.5/2.0 group should be similar to the 2.0/1.5 group performances at the very beginning as no

advice was given to either group. For 1.0/1.0 No-PFFB, this was not the case. At the start of training with no feedback given, average performance by flight number for the No-PFFB group was much worse than the PFFB group. This continued throughout the training phase.

The transition between flights 5 and 6 for the 2.0/1.5 flights showed that for one flight, the PFFB group did better than the No-PFFB group and then proceeded to do worse for the rest of the training. This could have been from a loss of confidence in what the instruments were showing the participant and in the instructions given from the instructor. For the No-PFFB group they only relied on the instruments and scripted instructions from the experimenter.

CHAPTER III

LANDING TASK EXPERIMENT

This chapter presents an experiment in which the effect of Out-of-the Window (OTW) visual cues for training of novice pilots for a straight in landing approach on a flight simulator was studied. A performance metric based on RMS errors in the landing flight parameters was used. To validate the chosen computed metric, the computed grades for a few selected landing flights were compared with grades determined independently by three certified flight instructors by viewing the flight video of the approaches.

Setup

Hardware

The Landing Study was performed in the Flight Vehicle Design Lab at Tuskegee University. The experimental setup consisted of one computer with 15" monitor used as the instrument panel or heads-down-display (HDD); three other computers with monitors were used for OTW. The physical setup, shown in Figure 9, is a mock setup of a partial cockpit.

The three OTW view computers are all Gateway E series computers with the following specifications: Pentium 4 2.60GHz processor Hyper-Threading Technology with 512KB cache, 512 MB ram, bus speed 800 MHz and memory speed of 333 MHz; NVIDIA's GeForce FX5200 is the video card used for the three OTW view computers. The "master" computer, which includes the HDD, is a Dell with following specifications: Pentium 4 2.4GHz processor with 512 KB cache, 512 MB ram, bus speed 133 MHz. The video card is a NVIDIA GeForce Ti 4600 with 128MB memory. A KVM (keyboard, video, and mouse) switch was used to provide keyboard and mouse service to all computers. Three 19"

monitors were used to provide the 135 degree out of the window (OTW) view Figure 9. The fourth monitor is a 15" monitor used to view the instrument panel. All of the four computers use 100 megabit Ethernet cards for LAN access. The computers are physically connected using a 5-port 100 megabit switch. CH rudder pedals and Thrustmaster Cougar HOTAS stick and throttle were used for flight and engine controls.

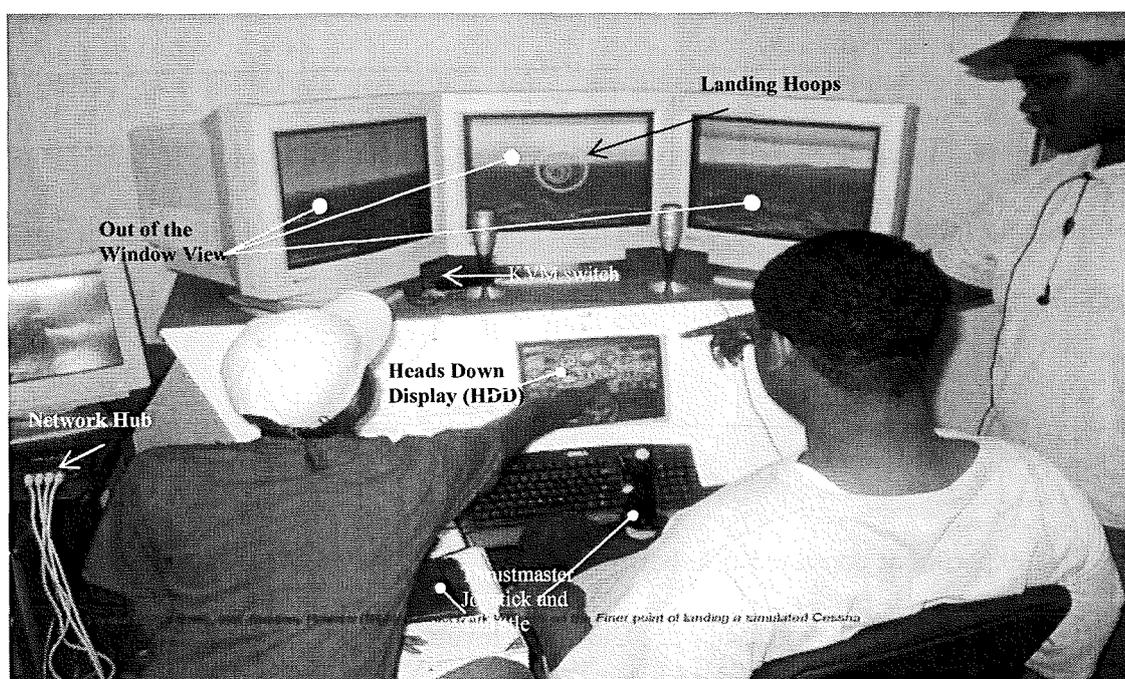


Figure 9. Partial mock cockpit. MSFS2002 shows the hub (switch) OTW view, HDD, switch and Thrustmaster joystick and throttle (partially hidden)

Software

The operating system for each computer is Microsoft Windows XP. Microsoft Flight Simulator (MSFS) 2002 was installed in the default location on each PC. In order to create the panoramic OTW view, Wideview's *.dll file was installed. Wideview¹⁷ is a program that will allow multiple instances of Flight Simulator to be controlled by a "master" computer so

that a panoramic view can be attained. Using Wideview, it would be conceivable to have a 360° display, given enough computers. As a prerequisite for Wideview's capability to synchronize OTW view displays, FSUIPC.dll and the IPX/SPX protocol (for Windows) were installed. FSUIPC¹⁸ is a pathway for programs like Wideview to talk to MSFS 2002. The flight recorder software module FLTREC.dll¹⁹ was installed to record the data to a *.dat file. FLTREC's output is configurable through a settings menu item. The data in the *.dat file is in XML format. A sample spread sheet was used to translate the fltrec.dat file into Microsoft Excel workbook columns for a spreadsheet. An additional worksheet was created to determine the performance (based on a 4.0 scale). The 4 point grade scale for straight and level flight was similar to that used by Williams¹⁷. The grade was based on values in certain ranges for parameters of level flight.

Method

Experimental Design

This experiment consisted of a between groups design with the independent variable being the type of visual cues presented. The three levels of the variable were: all training flights with OTW cues, all training flights without OTW cues, and a few training flights with OTW cues while others without OTW cues.

Participants

The student pilots participating in the experiment were all freshmen attending Tuskegee University. As an incentive, extra credit in an Introductory Psychology course was awarded to every participant. The volunteers that were selected had little or no prior experience with flight simulators or piloting an airplane. Every volunteer, after filling in

consent and background forms (Appendix C), attended a brief orientation session and conducted five familiarization flights. The orientation session introduced the various controls and instruments that the participant would use during the experiment. Three of the familiarization flights required the student to fly straight and level. Any volunteer who scored a 3.0 or better on at least one of the familiarization flights was selected for the training and evaluation sessions. The grading matrix is given below in Table 2. The excel Visual Basic module that calculated the grades is shown in Appendix A.

Throughout the experiment, every participant was individually attended by at least one of the three available experimenters. The experimenters were the author, and two other undergraduate senior Psychology students. None of the senior Psychology students had any experience on a flight simulator, so they spent a good deal of time becoming familiar with the equipment by asking questions and teaching each other the operating procedures.

Procedure

Orientation to Flight Controls and Aircraft Instruments

For the Orientation session, pilots were shown the relevant instruments: airspeed indicator, vertical speed indicator, artificial horizon, gyroscopic compass and RPM gage shown in Figure 10. Appendix D includes the set of instructions that were given to every volunteer in the orientation session. The pilot controls introduced were: Joystick, throttle, and rudder pedals. Instructions were given on how to operate the controls and the consequences of operating them. For example, 'pulling back on the joystick pitches the nose up'.

Blank Page

Familiarization Flights

Each participant flew 5 familiarization flights. The first three familiarization flights were straight and level flights of three minutes duration each. If necessary the experimenter could provide hands on help to steady the airplane in flight. During the straight and level familiarization flights, participants were graded on a 4.0 GPA scale. The scores were computed using a Visual Basic 6.0 code that compared the parameter values in flight recorded every second with pre-determined parameter values and tolerances. The pilot was required to fly 5,000 ft, 75 knots, 0 degree bank angle. For each flight a score was calculated based on the following criteria:

Table 2. Grading Criteria for a Straight and Level Flight

Nominal altitude = 5000 ft Nominal Speed = 75 knots Nominal Bank angle = 0 degrees			
Deviation(Δ in altitude) feet	Deviation(Δ in speed) knots	Deviation(Δ in bank angle) degrees	Grade
$\Delta = \pm 50$	$\Delta = \pm 5$	$\Delta = \pm 3$	A
$+50 \leq \Delta \leq +100$	$+5 \leq \Delta \leq +10$	$+3 \leq \Delta \leq +6$	B
$-100 \leq \Delta \leq -50$	$-10 \leq \Delta \leq -5$	$-6 \leq \Delta \leq -3$	
$+100 \leq \Delta \leq +150$	$\pm 10 \leq \Delta \leq \pm 15$	$+6 \leq \Delta \leq +9$	C
$-150 \leq \Delta \leq -100$	$-15 \leq \Delta \leq -10$	$-9 \leq \Delta \leq -6$	
$+150 \leq \Delta \leq +200$	$\pm 15 \leq \Delta \leq \pm 20$	$+9 \leq \Delta \leq +12$	D
$-200 \leq \Delta \leq -150$	$-20 \leq \Delta \leq -15$	$-12 \leq \Delta \leq -9$	
$+200 < \Delta$ or $\Delta < -200$	$+20 < \Delta$ or $\Delta < -20$	$+12 < \Delta$ or $\Delta < -12$	E

Those participants achieving a score of 3.0 or greater continued with the training. The fourth and fifth flights familiarized the pilot with making a straight-in landing initiated approximately from an altitude of 2860 ft with 500 ft/min rate of descent without any OTW

cues. In both flights, the experimenter provided verbal feedback to the pilot to assist in making appropriate corrections. In the fifth flight, the experimenter gave instructions to the pilot to listen to the voice from the simulator and where necessary, make appropriate corrections. After each of the familiarization flights, a graphical record was shown to the pilot using Microsoft's Flight Analysis Tool. This graphical record showed a trace of the altitude and heading of the aircraft.

Training Flights

The training session required every pilot to fly eight straight-in landing flights from an altitude of approximately 2860 ft and at a heading of 192 degrees thus aligned with the centerline of the runway. The task given to the pilots was to maintain 75 knots airspeed, descend at 500 feet per minute (by reading both the vertical speed indicator and by looking at the V.A.S.I (Visual Approach Slope Indicator) lights on the left side of the runway, and head towards and line up with the centerline in a straight-in approach to the Bremerton International Airport in the Seattle, Washington area. Each pilot was asked to land on the centerline of the runway. The aircraft being flown was a Cessna 172. The flight scenario was a modified version of a landing lesson from the "flight school" portion of Microsoft Flight Simulator, MSFS 2002. Participants selected for training were randomly assigned to the three groups named HH, NH and FH. A trainee in group HH was required to conduct all of the eight training flights with hoops as OTW cues. The flights with the hoops were a modified version of the 'Landing Lesson II' of the MSFS 2002. The hoops in this lesson are placed on the glide-slope. It was explained to the participants that if they maintained the required flight parameters, they would automatically pass through the hoops. For the NH

group a trainee conducted all of the flights without the OTW cues. A trainee in group FH was required to conduct eight flights some of the flights had hoops and the remaining flights had no hoops. The selected density was '5 hoops in sight at a time', which was the lowest setting in Microsoft Simulator.

Feedback in Training Flights

After each training flight, participants were able to look at the flight analysis provided by Microsoft Flight Simulator 2002. In the analysis the student could see a quick playback of a trace of the flight and see whether the airplane was on the centerline of the runway and whether it was descending at a constant rate. MSFS2002 also provided a computerized voice for training a landing task.

Evaluation Flights

After the training flight two evaluation flights were flown. The same scenario as the training flights was used. The only difference was that all trainees had to fly the glide slope without hoops. The sound was turned off so that the trainee could not hear the computer voice. Also, Microsoft's version of mild turbulence was added to make the task more challenging. No feedback was given after each of the flights. The total duration to complete the three phases for each participant was approximately 2-1/2 hours.

Observation

At the end of the experiment, it was noted that all participants in the FH group were given two distinctly different sequences of hoop flights. Therefore, the FH group was dropped from the study. However, the evaluation flights were used for the CFI study reported in Chapter IV.

Performance Metrics

Metric Definition

Data were gathered using third party flight recorder software Fltrec9.0 software. This software writes its data to an .xml file and was interpreted by a shareware spreadsheet into MSExcel for each flight. The analysis originally included data from the time the simulator was un-paused to the time the pilot should have touched down. This was not really accurate in that sometimes before the runway threshold, pilots must reduce power; change the airplane rate of descent (and airspeed) to make a landing. We decided to calculate the time it should have taken an airplane descending at 500 ft/min to hit the ground (which is what would happen) and compare this to the actual flight. Most of the flights exceeded this time.

The following parameters were taken into consideration for determining the landing task performance metric: rate of descent, runway alignment, and airspeed. Thus error in speed is $(\Delta V = V_r - V_f(t))$, and error in rate of descent is: $(\Delta ROD = H_f(t) - H_r(t))$

where
$$H_r = H_{start} - 8.333 \text{ft/sec} * t$$

The performance metric used assumed the airplane was on the extended centerline of the runway. For each second an airplane has a heading and speed. Then for each second, the airplane will have a distance Δx that it will move laterally from or toward the centerline. The distance the airplane travels, in a straight line, is: $= V * t$.

Error in the heading, $(\Delta \psi)$ of the airplane is the difference between the runway heading and the current heading.

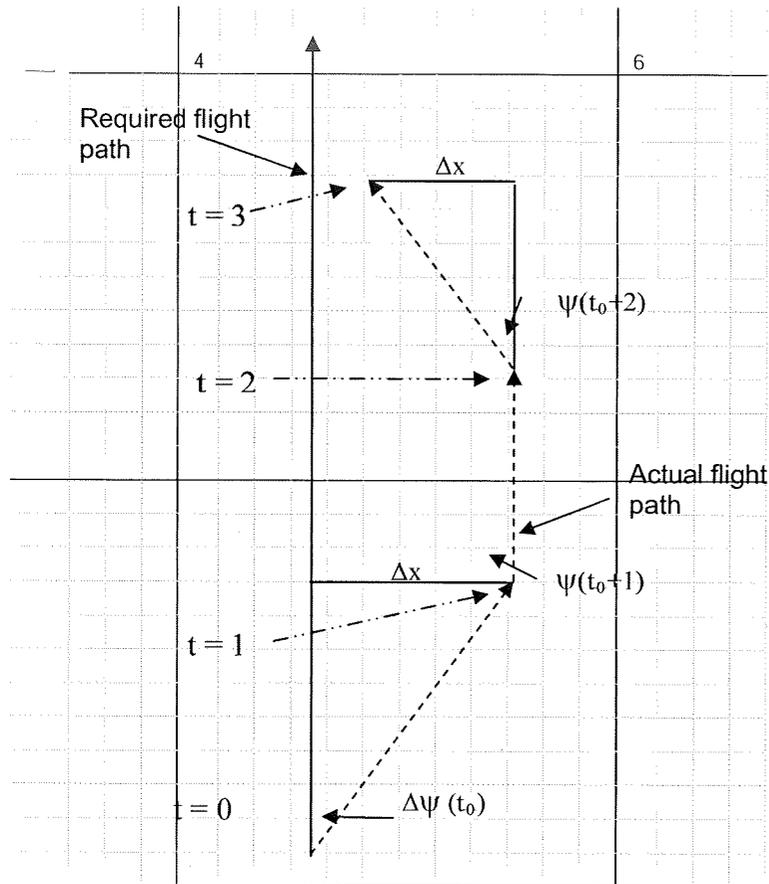


Figure 11. Runway Alignment.

The straight-line distance and the $\Delta\psi$ heading, lead to the distance

$$\Delta x(t) = V(t) * t * \sin(\Delta\psi(t-1))$$

Error in heading ($\Delta\psi = \psi(t_0) - \psi(t_0+1)$), $\psi(t)$ heading at time t is in radians

Since velocity is in knots conversion to feet was necessary

$$\Delta V \text{ ft/s} = 1.687809 * (\Delta V \text{ knots})$$

Variables Lat1 through Lat5 (Figure 12) were Degrees from the Prime Meridian and converted to radians for calculation. Error in each of these parameters was determined for

every second of the flight by taking the difference between the required value of the parameter and the value of the parameter during the flight.

V_r is the required velocity

$V_f(t)$ is the velocity at time t

ψ_r is the heading in radians

$\psi_f(t)$ is the heading at time t

H_r is the required altitude

$H_f(t)$ is the altitude at a time t .

H_{start} is the starting altitude

Δx is the lateral error from centerline

$H_f(t)$ is the altitude at time t assuming a descent rate of 8.333 ft/s

The errors, ΔV , Δx , and ΔROD are calculated for each second of a flight which could vary from 300 – 420 seconds. The RMS errors for ΔV , Δx , and ΔROD were computed for the entire flight. A linear combination of the RMS errors was used as the metric for evaluating the performance of the segment of the flight where the airplane should have constant airspeed, rate of descent and heading.

Starting Error

The airplane however, did not start out on the centerline. So each participant had to align the airplane and descend at the proper rate and the correct speed. This slight offset was barely discernable on the monitor but very noticeable in the data. To identify this problem, Latitude and Longitude data were gathered and used as “Cartesian coordinates.” The straight line distance between the current Lat/Long and the centerline Lat/Long would then be the Δx . Using this method it was determined that the flight started out about a third of a mile 1600 ft left of the centerline.

The method used to calculate the distance from the centerline Δx starts with the figure shown in Figure 12. The latitude and longitude for the point on the centerline was

determined with the method below. The straight line distance in kilometers from the centerline (Δx) was determined using equation 6 below¹⁵, that comes from the Spherical Law of Cosines¹⁵ and using 6371km as Earth's radius (R).

$$d = \text{acos}(\sin(\text{lat}_1) \cdot \sin(\text{lat}_2) + \cos(\text{lat}_1) \cdot \cos(\text{lat}_2) \cdot \cos(\text{long}_2 - \text{long}_1)) \cdot R$$

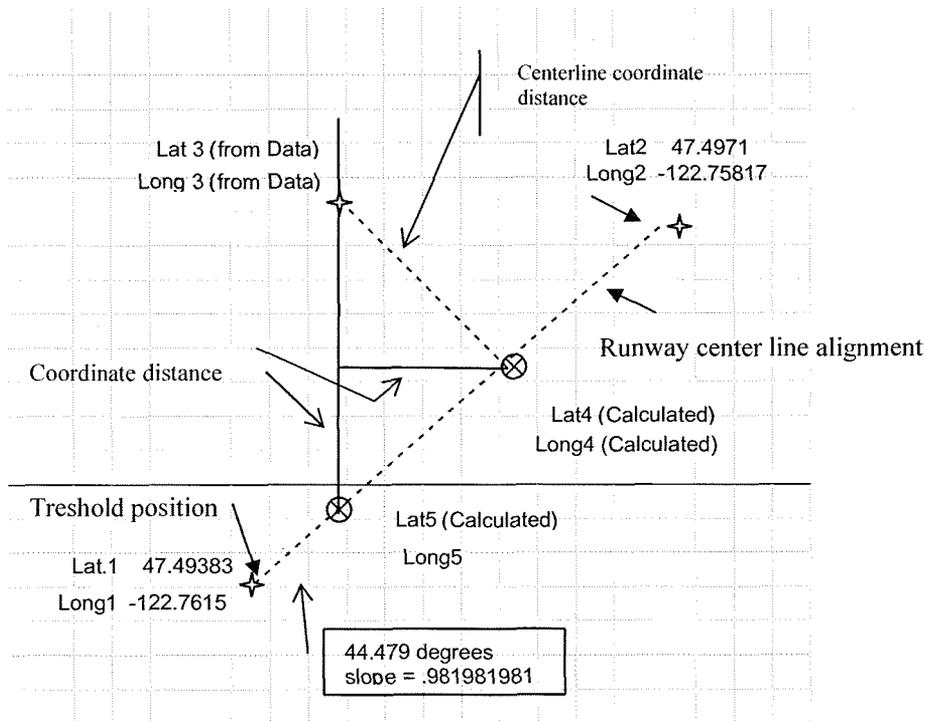


Figure 12. Runway alignment error geometry.

$$\text{Lat5} = .981981(\text{Long3}) + 168.043 \quad \text{-----}(\text{Eq. 3.1})$$

$$\text{Centerline_coordinate_distance} = (\text{Lat3} - \text{Lat5})\text{Cos}(\tan^{-1}(.981981981))\text{-----}(\text{Eq. 3.2})$$

$$\text{longitude_coordinate_distance} = \text{Centerline_coordinate_distance} * \text{Cos}(\tan^{-1}(-1/.981981981))$$

----- (Eq. 3.3)

$$\text{Long4} = \text{Long3} + \text{longitude_coordinate_distance} \quad \text{-----} (\text{Eq. 3.4})$$

$$\text{Lat4} = .981981(\text{Long4}) + 168.043 \text{-----}(\text{Eq. 3.5})$$

Alignment Error

$$\Delta x \text{ (km)} = \text{ArcCos}(\sin(\text{Lat3}) * \sin(\text{Lat4}) + \text{Cos}(\text{Lat3}) * \text{Cos}(\text{Lat4}) * \text{Cos}(\text{Long3} - \text{Long4})) * 6371$$

(ref. 15) --- (Eq. 3.6)

$$\Delta x \text{ (ft)} = \Delta x \text{ (km)} * 1000 \text{ meter/km} * 3.2808339 \text{ ft/meter} \text{-----} \text{ (Eq. 3.7)}$$

The flight data were processed to determine the error in airspeed, rate of descent and runway alignment. A linear combination of the non-dimensionalized standard deviations of errors in the three parameters was then calculated and used as a composite error metric for a flight as well.

Results and Discussion

The performance metric was plotted against flight number for the purpose of determining whether student pilot's performance improved over time. In this performance metric, performance improvement means a decrease in the value of the metric. Figure 14 presents the performance metric, normalized with the first training performance, for every participant in the No Hoop (NH) group and for every flight in the training and testing sessions. Likewise, Figure 16 presents the performance metric, normalized with the first training performance, for every participant in the HH group. Figure 17 presents the performance metric of every flight in the training and testing sessions for the NH and HH groups separately averaged over the 12 participants per group. A simultaneous study of Figures 14, 16, 17 leads to the following observations.

In both NH and HH groups, most of the pilots have improved performance in the second training flight (T2) when compared with the first training flight (T1). In any typical

training program it is expected that the performance of trainees would show a visible improvement with the amount of training; surprisingly, such a trend is not seen here for the six training flights in sequence from T3 to T8. For the HH group, the average performance for the last five flights is visibly better than that of the first three flights. But the average is unduly influenced by the performance of participant #25. One of the participants in the NH group, #31, offers a strange pattern of alternating increase and decrease of performance from T1 to T8. Few other participants in both groups show a visible decrease in their performance at different stages of their training sessions. Also, there seems to be no improvement in performance for both HH and NH participants from T6 – T8. This could mean that perhaps our training task is too easy, and the number of flights is too many so that the participants have reached a plateau in their training and can't be further trained in that scenario.

The transition from training to evaluation going from T8 to E1 shows that most of the participants in both groups have performed worse in the first evaluation flight than in the last training flight. That is understandable due to a relatively more challenging task in the evaluation flights when compared with the training flights. For the trainees in the NH group the added challenge in the evaluation flight is the presence of mild turbulence. For the trainees in the HH group, the absence of hoops in the evaluation flight is another challenge together with the presence of turbulence. Surprisingly, in the transition from training to evaluation, the additional challenge for the HH group did not result in a larger decrease in performance than that in the NH group. This could be taken as a positive aspect of training with hoops as compared with training without hoops.

Figure 17 shows that the performance of the NH Group, on average, was worse than the HH group when transitioning from the last training flight to the first evaluation flight. The hypothesis was that a training intervention, such as tunnel in the sky, would be a better way to train pilots than would standard methods without a training intervention. Some evidence of this may be seen in the first evaluation flight (E1) in which performance of the HH Group was slightly better than the NH Group (Figure 17).

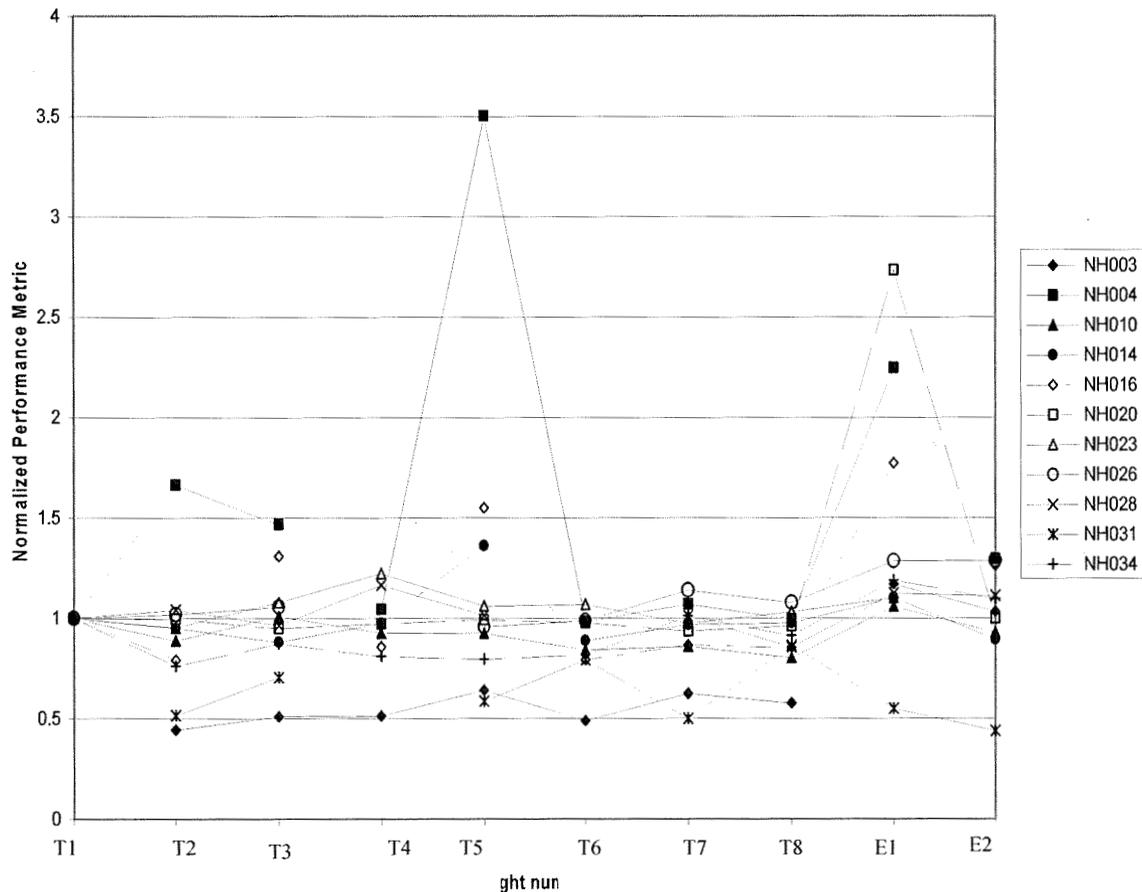


Figure 13. Normalized performance of the participants in NH group.

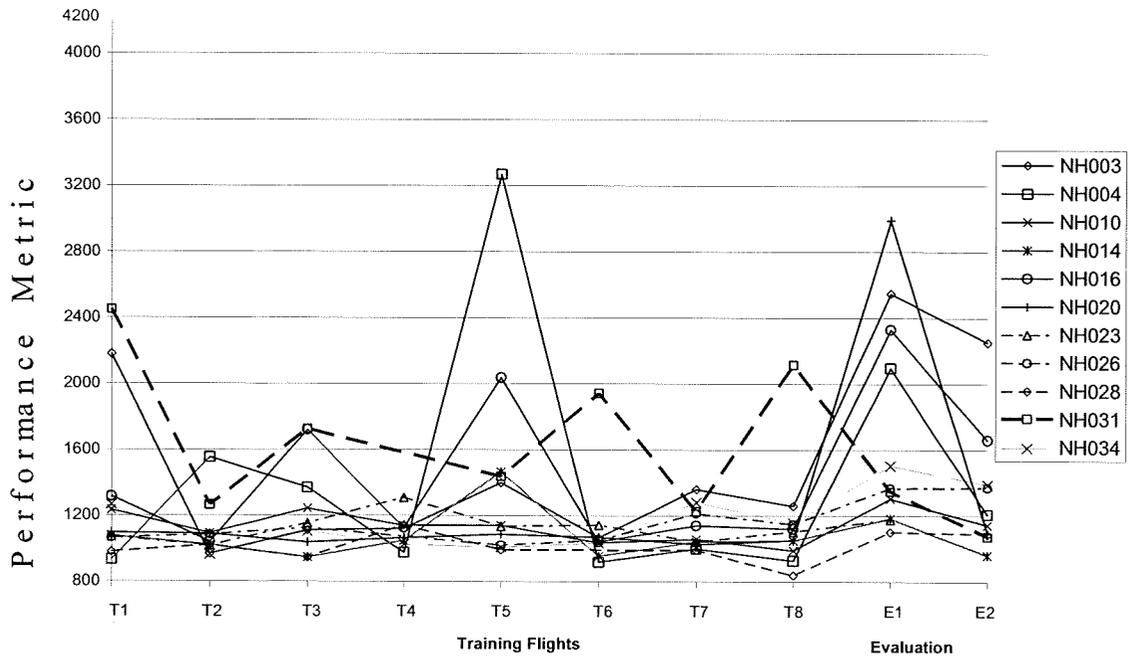


Figure 14. Performances of the participants in the NH group.

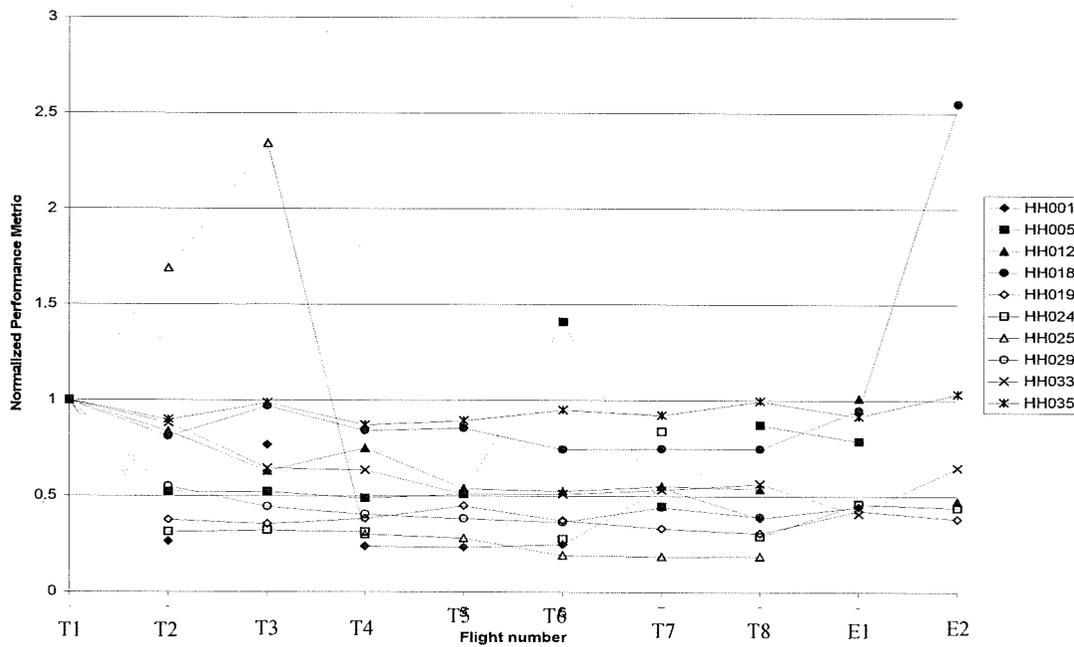


Figure 15. Normalized performance of the participants in the HH group.

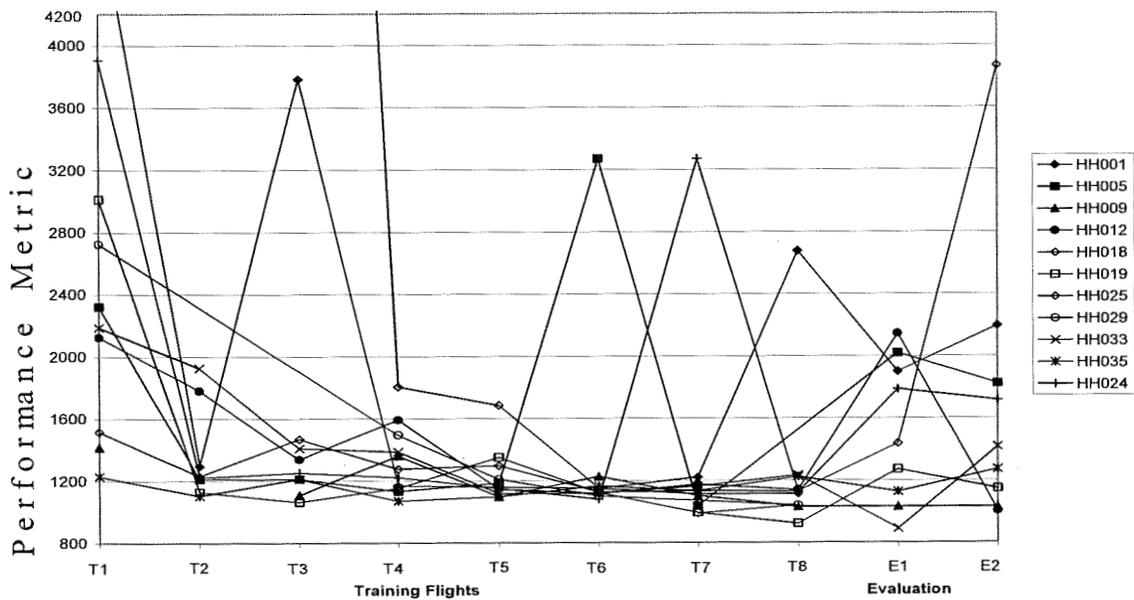


Figure 16. Performance of the twelve participants in the hoop (HH) group.
 (Missing data: T5 for HH001, T8 for HH005, T5 HH024, (E1, E2) from HH025, (T2, T3) for HH029, HH009 data for T8, E1, E2 is suspect; Not shown here T1 for HH001; T1-T3 for HH025)

Average Performance Metric for Hoop and No Hoop flights

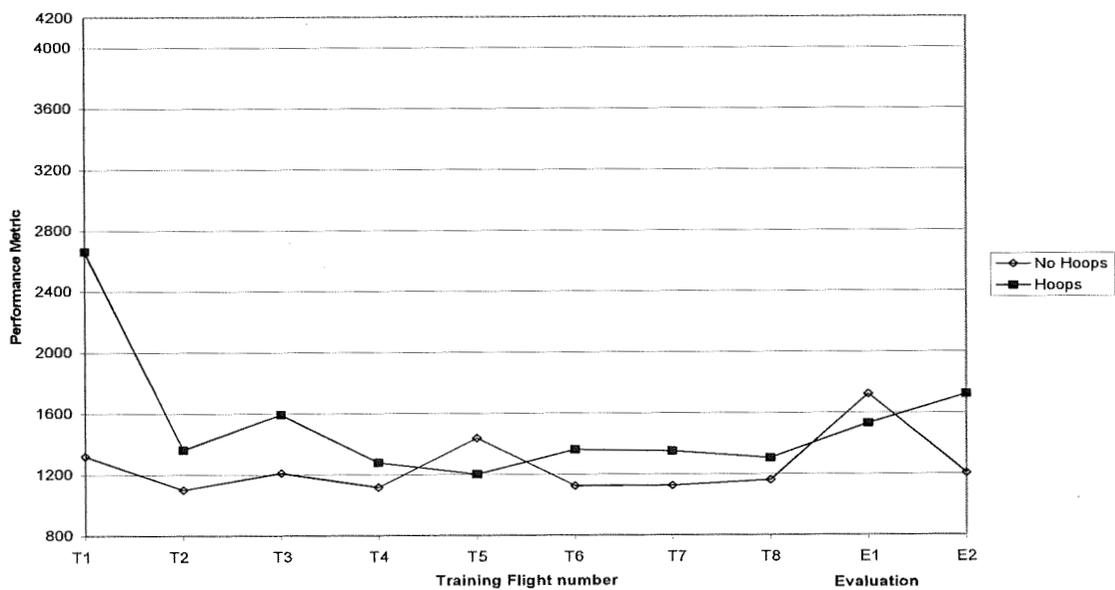


Figure 17. Average performance of the NH and HH groups.

CHAPTER IV

COMPARSION OF COMPUTED METRIC WITH CFI EVALUATION

Introduction

For the student pilot in a FAR Part 61 or FAR Part 141 flight school, ten flight hours of required training can be flown in a PCATD. However, the regulations state that a CFI must also be present to supervise the training (FAA 1997). Vruels and Obermayer¹³ emphasize the need of automated performance measures on modern simulators as a substitute for direct evaluation of performance by an instructor. To address this goal for training of a landing maneuver, the present effort included a study of how closely a computer can emulate a CFI's evaluation of a student pilot.

Method

As mentioned above, the study of landing performance of novice pilot on MSFS 2002 required all trainees to fly two evaluation flights under identical conditions irrespective of the strategy used in their training flights. These landing approaches were video recorded using the MSFS2002 capability so that the flights could be independently evaluated by three CFI's with a minimum of 5000 hours of experience (for details see Appendix E).

The CFI's were required only to grade the approach, and not the flare and touchdown. CFI's were not informed regarding the training intervention experienced by the pilots. They were asked to assign each flight a grade based on a 100-point scale. They were also requested to monitor three flight parameters namely, airspeed, rate of descent and runway alignment. Of those parameters they were to indicate the best and the worst controlled flight

parameters and the flight parameter most affecting the grading. The details of the instructions to the CFIs are given in Appendix D.

Results and Discussion

For the thirty-three evaluation flights graded individually by the three CFI's, correlations between the pairs of instructors' grades were found to be rather low ($r^2 = 0.43$, 0.51 , and $.67$; see Figure 18.) Therefore, the CFI scores were further analyzed for flights in which the CFIs agreed on both the best and the worst controlled parameters for a flight.

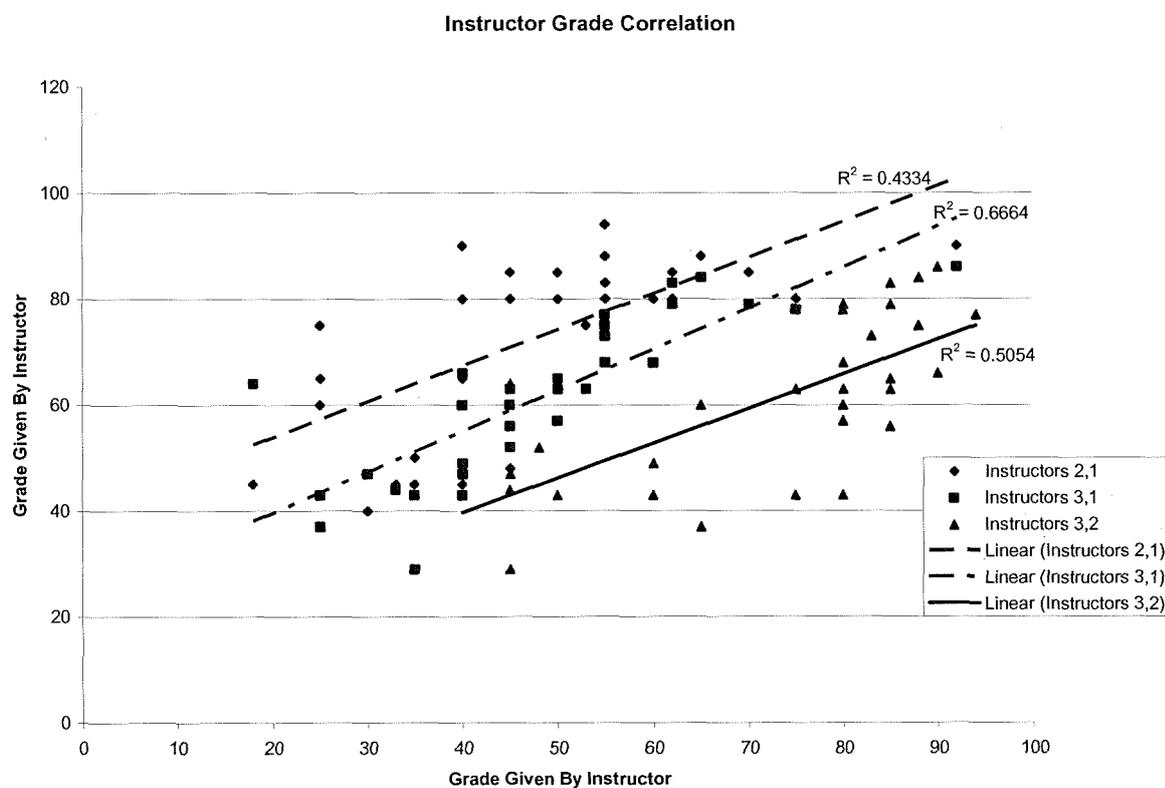


Figure 18. Correlation of instructor grades.

All three CFIs cited runway alignment as the best controlled parameter more often than the other two parameters (Figure 19). Also, rate of descent was identified as the worst controlled parameter more often by all three CFIs (Figure 20). All three instructors agreed on both the best and worst parameters for seven flights. For grading of the seven identified flights, despite their agreements, a high correlation was seen only between the grades of two CFIs ($r^2=0.89$, Figure 21). The grades of the third CFI had low correlations with the grades awarded by either one of the other two CFIs ($r^2= 0.44$ and 0.25 , Figure 21).

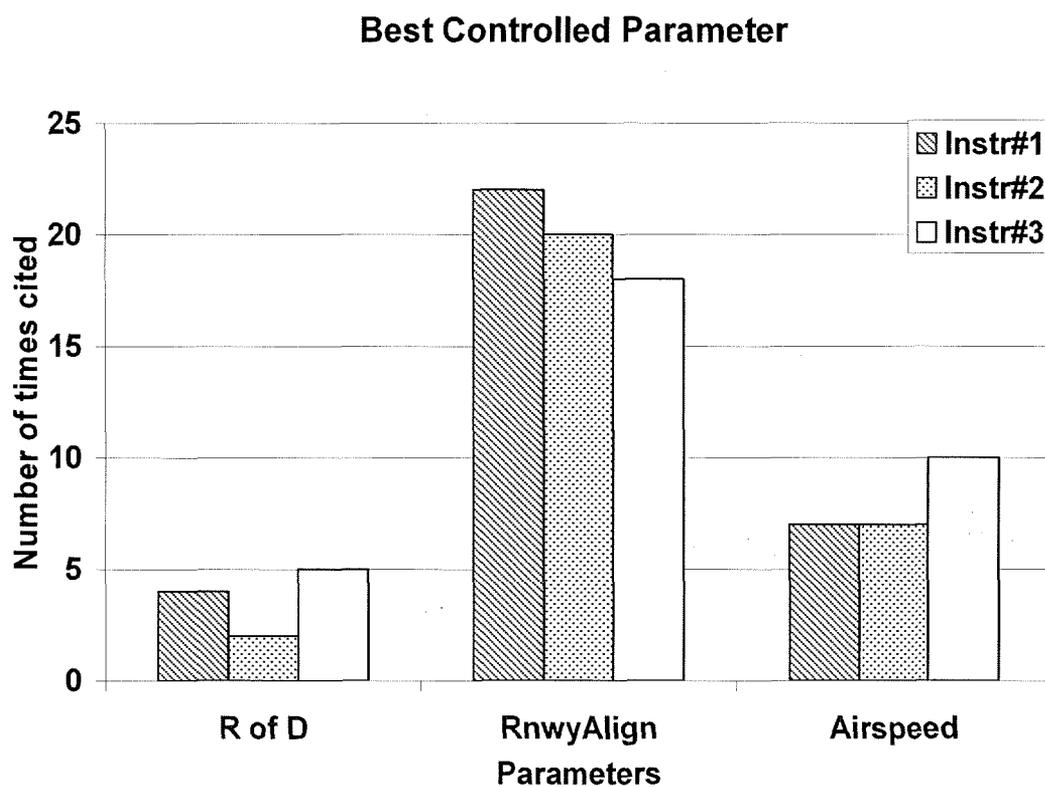


Figure 19. Best controlled parameters as cited by the CFIs.

Worst Controlled Parameter

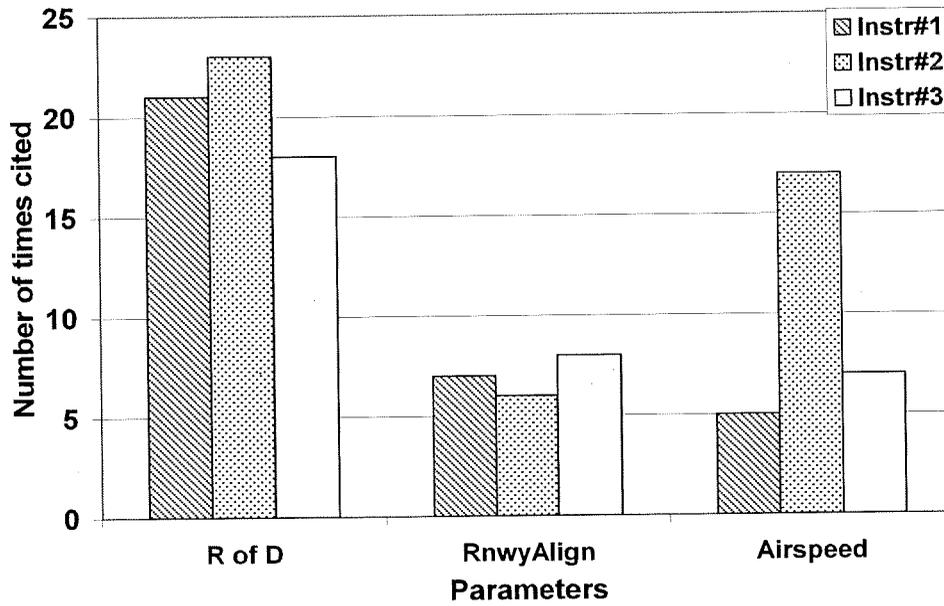


Figure 20. Worst controlled parameters as cited by the CFIs.

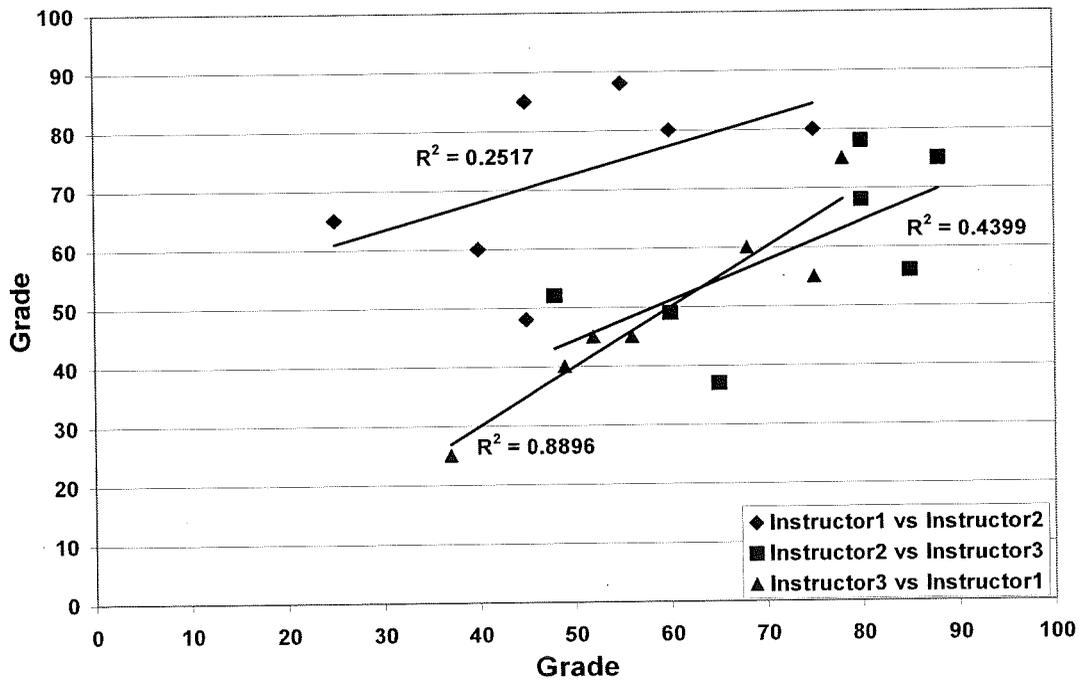


Figure 21. Correlations between flight instructor scores of selected flights.

For the seven landing flights identified, the calculated composite error metric values were plotted on a graph separately against the grades awarded by every one of the three instructors as shown in Figure 22. For the composite error metric, strong and significant correlations were found with the grades awarded by every one of the three instructors ($r^2 = 0.82$, $t = 4.69$, $p < 0.01$; $r^2 = 0.77$, $t = 4.15$, $p < 0.01$; and $r^2 = 0.63$, $t = 2.91$, $p < 0.05$).

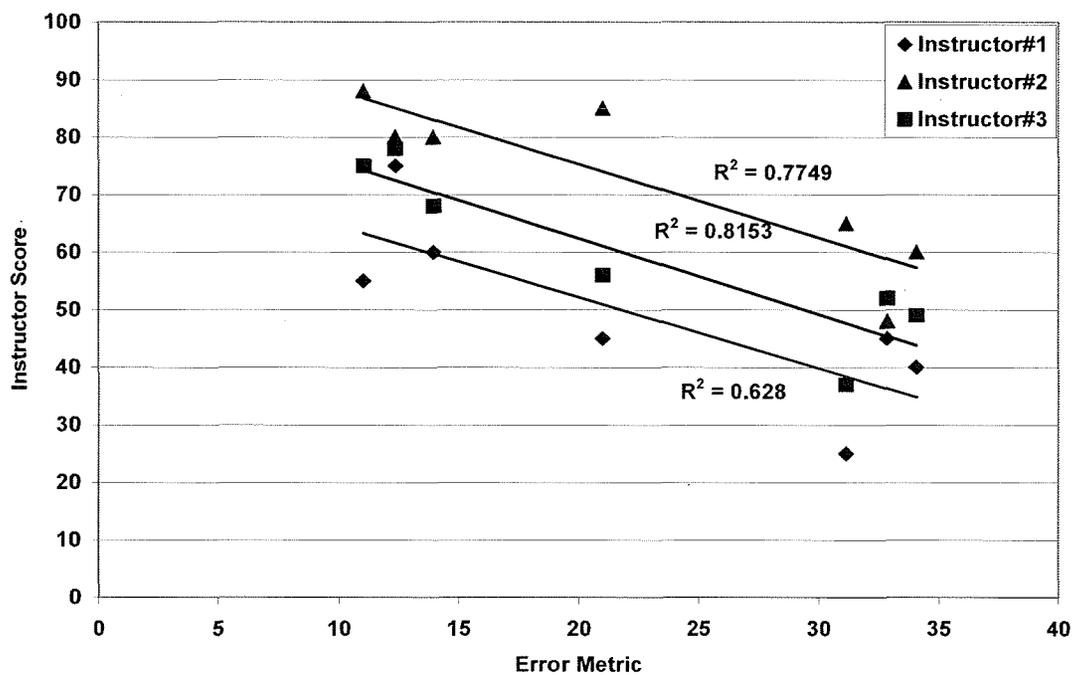


Figure 22. Correlation between composite error metric and instructor scores.

These correlations indicate the suitability of the composite error metric to be considered for use as a measure of performance on a straight-in landing approach. The fact that correlations between instructors on grading a landing approach were low, and that such grades correlated poorly with the performance metric based on RMSE in the flight parameters of rate of descent and runway alignment indicate the difficulties inherent in this

research area. The strategy of determining first the flights in which instructors agree on best and worst controlled parameters is a promising one. Such a strategy may be effective in narrowing the focus to fewer variables than what would normally be involved in grading a flight maneuver, and should lead to a better understanding of the variables involved in CFI evaluation of a maneuver. This strategy could then be applied to different maneuvers and perhaps weights could be applied differentially to different parameters depending upon the parameters most important in evaluation of that maneuver.

CHAPTER V

CONCLUSIONS AND SUMMARY

This thesis examined different training interventions for novice pilots in an attempt to develop techniques and performance measures that could be incorporated into an ‘intelligent’ training system. Effectiveness of training interventions were investigated by comparing performance on a level turn and straight-in approach maneuver flown by novices on a flight simulator with control groups flying the same maneuvers on the flight simulator. Metrics were defined to compute performance using flight parameter data of the maneuvers. Performance computed from flight parameter data of the treatment-group and control group flying the straight-in approach was also compared with certified flight instructor evaluation of the video data of the same flights to validate the metric used for computing the performance.

The following conclusions are drawn from these studies:

- Real Time Training (RT) with post flight feedback resulted in the best performance on the evaluation flight for a level turn as compared to no feedback or training with ARTT.
- Above Real Time Training (ARTT) using a sequence of 1.5/2.0 with post flight feedback was observed to be the next best strategy for a level turn. Thus, if time is of the essence then this training strategy is a promising approach. Further studies with more participants could validate this finding and determine whether 2.0 for all flights, or 1.5 for all flights might be a more effective strategy.

- The particular post flight feedback provided to the participants flying the level turn, consisting primarily of graphical feedback which showed the actual ground track superimposed over the target ground track may be an effective form of feedback to implement in an "intelligent" flight simulator providing automated feedback.
- For the landing task, no statistically significant difference was found between the group provided with visual out-of-the-window cues of hoops and the control group. This indicates that the additional visual cues did not add to the training value, perhaps because visual cues in the form of the runway and visual approach slope indicator lights were already available for this maneuver.
- Poor correlations were observed between inter-CFI evaluations. Where the CFIs agreed on the best and worst controlled parameters, correlations were obtained between the computed performances using the performance metric developed in this study and the evaluations of the CFIs for the straight-in landing approach task.

Future Studies

- The two studies in this volume were meant to be steps on the way to creating an intelligent flight simulator for training novices. To that end, it would be interesting to look at the following:
 - To validate any of these training strategies, however, a) more participants need to be studied to determine if the effects observed are generalizable and robust, and b) studies assessing whether these methods produce better performance in the actual aircraft need to be undertaken.

- Conduct more studies to determine the reliability and robustness of the effect of ARTT as a strategy for training. Specifically, compare RTT with ARTT 1.5 on the turn maneuver, both using feedback.
- Use additional feedback mechanisms to enhance the already provided graphical feedback, such as automated audio to correct for error.
- A study of a combination of tunnel in the sky and ARTT as a training intervention might yield interesting information.
- Each training intervention needs to be evaluated for more than one maneuver to determine its robustness.

APPENDICES

APPENDIX A

Visual Basic Scoring Module:

Function to strip repeating data caused by pressing pause button. Also calculates the overall score for the familiarization flight.

```

Function CalcScore()
Dim TotalScore As Double
Dim Count As Integer
Dim Stuff
' Module created by Bruce Heath
' December 31, 2003
,
' Module finds the beginning of relevent data and
' then adds the scores for 180 rows (seconds)
' The totaled scores are then averaged by 180 for
' the average score
,
' Version 2 February 6, 2004
' Complete February 12, 2004
' Corrects the problem caused by pausing the simulator for more
' than 300 seconds while the flight recorder is running

row = 2 'First row of data
col = 4 'Seconds column for Clock time
'MsgBox Sheets("Data").Cells(3, 4).Value

'This loop finds the starting point for the data

Do Until (Sheets("Data").Cells(row, col).Value <> Sheets("Data").Cells(row + 1, col).Value)
    row = row + 1
Loop

' Modification : If the row on the DATA sheet still has not reached the beginning of the
' data after end of the fomulated cells on the rows on the SCORE sheet,
' then find the beginning of the data on DATA and place the processed data on the SCORE sheet.
' Continue to place processed data on the SCORE sheet until the DATA sheet meets an empty
' row.
,
If ((Sheets("Score").Cells(row + 1, col).Value = "") And (Sheets("Data").Cells(row + 1, col).Value <> ""))
Then
    row2 = row
Do While (Sheets("Data").Cells(row2 + 1, col).Value <> "")
'Fill in Data

'Altitude
Sheets("Score").Cells(row2, 1).Value = Sheets("Data").Cells(row2, 9).Value - 3000

'Bank
If (Sheets("Data").Cells(row2, 10).Value >= 180 And Sheets("Data").Cells(row2, 10).Value <= 360)
Then
    Sheets("Score").Cells(row2, 2).Value = 360 - Sheets("Data").Cells(row2, 10).Value

```

```

Else
  Sheets("Score").Cells(row2, 2).Value = Sheets("Data").Cells(row2, 10).Value
End If

'Airspeed
Sheets("Score").Cells(row2, 3).Value = Sheets("Data").Cells(row2, 10).Value - 75

'Altitude Raw Score
With Sheets("Score")
  If (.Cells(row2, 1).Value < 0) Then
    .Cells(row2, 4).Value = (.Cells(row2, 1).Value / 50) * -1
  Else
    .Cells(row2, 4).Value = .Cells(row2, 1).Value / 50
  End If

'Bank Raw Score
.Cells(row2, 5).Value = .Cells(row2, 2).Value / 3

'Airspeed Raw Score
If (.Cells(row2, 3) < 0) Then
  .Cells(row2, 6).Value = (.Cells(row2, 3) * -1) / 5
Else
  .Cells(row2, 6).Value = .Cells(row2, 3) / 5
End If

Count = 1
Do While (Count <= 3)

'Assign Altitude Score
'Assign Bank Angle Score
'Assign Airspeed Score

Select Case .Cells(row2, Count + 3).Value
  Case Is <= 1
    .Cells(row2, Count + 6).Value = 4
  Case Is <= 2
    .Cells(row2, Count + 6).Value = 3
  Case Is <= 3
    .Cells(row2, Count + 6).Value = 2
  Case Is <= 4
    .Cells(row2, Count + 6).Value = 1
  Case Else
    .Cells(row2, Count + 6).Value = 0
End Select
  Count = Count + 1
Loop

'Three Score total for the row
.Cells(row2, 10).Value = .Cells(row2, 7).Value + .Cells(row2, 8).Value + .Cells(row2, 9).Value
'Average Score for the row
.Cells(row2, 11).Value = .Cells(row2, 10).Value / 3
End With
row2 = row2 + 1

```

```
Loop
End If

' Now find the numeric score
loopEnd = row + 181 'Loop for the three minute time period

scorecol = 11 'Column K average for each second on every row
'Count off 180 seconds totaling the average scores
displayrow = row

Stuff = MsgBox(displayrow, , "Data Begins in Row")

With Sheets("Score")
Do Until row = loopEnd
    TotalScore = TotalScore + .Cells(row, scorecol).Value
    row = row + 1
Loop

'The numeric score is this
.Cells(7, 14).Value = TotalScore / 180
'MsgBox TotalScore
End With
End Function
```

APPENDIX B

Informed Consent Form

INFORMED CONSENT FORM

DEVELOPMENT AND ASSESSMENT OF A NOVEL TRAINING PACKAGE USING SELF INSTRUCTION METHODS AND ABOVE REAL TIME TRAINING (ARTT) FOR BASIC MANEUVERING TASKS ON A FLIGHT SIMULATOR

As an undergraduate student of Tuskegee University, you are invited to participate as a trainee in an experimental research study. The study will develop and evaluate a new program of training on a flight simulator.

Your participation will require approximately two hours of flying on a simulator one day. You will take a seat as a pilot in a mockup cockpit located in the Flight Vehicle Design Laboratory on Tuskegee's Campus. You will control the flight movements of the joystick, throttle and rudder pedals while watching the flight parameter changes on the simulated instrument panel Heads Up Display (HUD) and the moving icon of the airplane on the computer screens. The available instructor will give you a set of instructions, allow you some time for familiarization and for asking questions, and provide feedback on your performance whenever necessary.

No physical or medical testing risks are involved in the experiment. The experiment, however, is not an official pilot training. As a possible risk, some of the simulator flight techniques that you will learn may not be applicable in real flying. Moreover, as it happens in any learning exercise, you may need to overcome temporary disappointments on possible lack of progress in performance, especially when your flight would face a simulated crash.

In general, the training will be an enjoyment and a learning experience. The data provided by you on a survey form and the data on your performance in the training will be coded to protect your confidentiality. Your participation is voluntary and you may withdraw from the experiment at any time. Your participation is voluntary and you may withdraw from the experiment at any time. Your participation according to your agreed upon schedule, however will be greatly appreciated by the investigators.

The successful completion of this research program will be a valuable contribution by Tuskegee University in improving the methods of training pilots on a simulator.

For any questions or concerns, please feel free to contact the investigators on campus: Dr. Syed Ali at 727 – 8853, Dr. Marcia Rossi at 727 – 8830 or Dr. Muhammad Khan at 727 – 8637. In the event of an unresolved grievance, you may also contact Dr. Stephen Sodeke, the Chairperson of the Human Subjects Review Committee (HRSC) at 727 – 8363.

The available investigator will give you a copy of this form to keep. By signing below, you are agreeing to participate in this study.

Signature of Participant/ Date

Signature of Investigator/ Date

Participant Demographic and Data Form Turn Study

Background Survey- Turn Study

Code: _____

Major: _____

Sex: _____ Male
 _____ Female

Age: _____

Are you: _____ Left Handed
 _____ Right Handed

Do you have any flight experience?
 _____ Yes
 _____ No

If so how many hours (estimate)? _____

How long ago was your most recent flight training, if any? _____

How many hours a week do you engage in video/ computer games?
 _____ light: 0 – 5 hours
 _____ medium: 6 – 13 hours
 _____ heavy: 13 or more hours

What type of video / computer games do you play?

_____ sports
 _____ fighting
 _____ war
 _____ cards
 _____ flight
 _____ mystery
 _____ other:

If so what type? _____

What type of controllers do you use?

_____ joystick
 _____ control pad
 _____ keyboard
 _____ arcade

Debriefing Form for Turn Study

Subject Code: _____

Ask them if they have any questions about what they were doing?

1) Questions asked by participant and answers

2) Would you describe what you did as (circle one)

very fun
 somewhat fun
 neutral
 somewhat boring
 very boring

3) Would you agree that the pre training instructions (before any flight) prepared you for the flight training sessions?

strongly disagree
 disagree
 can't decide
 agree
 strongly agree

4) Would you describe the training sessions as (circle one)

very difficult
 difficult
 neutral
 easy
 very easy

5) Would you agree that the flight tasks (the first flying sets) prepared you for the flight test?

strongly disagree
 disagree
 can't decide/ don't know
 agree
 strongly agree

Tell them the purpose of the study, i.e., to determine if training at different ART levels for turning is more effective than RT training. Answer any questions. Ask them not to reveal the nature of the experiment to any others, as it may disrupt their performance. Include any comments they made that seem relevant, that might help in explaining their performance.

APPENDIX C
Participant Demographic, Data Form, Debriefing Form, Landing Study
 Background Survey- Landing Study

Code: _____

Major: _____

Sex: _____ Male
 _____ Female

Age: _____

Are you: _____ Left Handed
 _____ Right Handed

Do you have any flight experience?
 _____ Yes
 _____ No

If so how many hours (estimated)? _____

How long ago was your most recent flight training, if any? _____

Do you have any flight simulator experience? Yes _____ No _____

Overall, how much exposure have you had to Microsoft Flight or Combat Simulator? _____

How many hours a week do you engage in video/ computer games?

_____ Light: 0-5 hours
 _____ Medium: 6-13 hours
 _____ Heavy: 13 or more

What type of video / computer games do you play?

_____ Sports
 _____ Fighting
 _____ War
 _____ Cards
 _____ Flights
 _____ Other:

What type of controllers do you use?

_____ Joystick
 _____ Control Pad
 _____ Keyboard
 _____ Arcade

Have you ever belonged to an organization similar to the C.A.P. or Redtails?

_____ Yes If so what organization _____?
 _____ No

Are you familiar with any of the following concepts, if so please check?

_____ Throttle _____ Aileron _____ Rudder

Blank Page

Debriefing For Landing Study

Participant Code: _____

1. **Would you describe what you did as (circle one)**
Very fun
Somewhat fun
Neutral
Somewhat boring
Very Boring
2. **Would you agree that the verbal instructions and the orientation instructions (before any flight) prepared you for the training flights?**
Strongly disagree
Disagree
Can't decide
Agree
Strongly Agree
3. **Would you describe the training sessions as (circle one)**
Very difficult
Difficult
Neutral
Easy
Very easy
4. **Would you agree that the training sessions prepared you for the flight tests (the last two flights without sound)?**
Strongly disagree
Disagree
Can't decide
Agree
Strongly Agree
5. **Are there any questions about what you were doing?**

The purpose of the study: To determine if the use of cues (hoops and automated pilot) was effective for landing tasks. Please do not to reveal the nature of the experiment to any others, as it may disrupt their future performance. Thank you for your participation.

APPENDIX D

Instructions and Data form for Flight Instructors

Thank you for your participation in the evaluation. Each of these flights was made by an inexperienced student. Most of the students had no experience flying an airplane and, with few exceptions, no experience with simulators. What we would like for you to do is to view each flight (instructions below) and give the flight a grade. I realize that flight instructors usually do not give a grade on a point scale. However, for these flights please give the overall flight a grade based on a 100 point scale. After each flight, we would like for you to record the parameter which most influenced your grade choice. Also, we would like for you to identify the parameter the student best controlled, and the parameter the student controlled the worst (Runway alignment, glideslope, or airspeed). The experimenters are looking at three instruments: heading, vertical speed indicator and airspeed.

Instructions for Evaluating Flights

We would like for you to evaluate these flights for a straight-in approach to landing. Each approach was made with turbulence turned on. They should **not** be graded for the flare and touchdown.

For the following flights, the student participants were asked to perform a straight-in approach to landing and to maintain alignment with the runway, a descent rate of 500 fpm, and 75 knots. Please note that the approach was being made with mild turbulence.

For these flights, please make sure to turn off the sound as the students could not hear the instructor when they made the approach. Also, although messages occasionally appear on the top left corner of the screen from the automated instructor, the students were not given any instructions regarding responding to them. Therefore, please disregard those messages when conducting your evaluation. Please begin your evaluation after the AP light goes off (the AP light is on the top left side of the instrument panel.)

Thank you,
Bruce Heath

To load the video files of the flights please use the following directions:

We would like for you to evaluate flights for our participants. Microsoft Flight simulator's video recorder was used to make a visual record of each of the flights. In order to replay the video which displays the instruments and a forward outside view, please use the following directions:

Copy all of the files from the CD to :

C:\Program Files\Microsoft Games\FS2002\flights\myflts

Start Microsoft Flight Simulator 2002

At the opening screen select "Learn to Fly" from the menu on the left.

Under "2. Choose a Lesson" select **Lesson 6h: Landings Part II**

Click the **Fly Now** button in the lower right corner.

On the next screen click **OK**.

As soon as the scenario starts, hit the **ALT** key on your keyboard.

Select **OPTIONS** from the menu that should appear at the top of the screen.

Click **Flight Videos** from the **OPTIONS** menu.

Select one of the flights under **SAVED VIDEOS** and click **PLAY CLIP**.

The simulation will start paused. Hit the **P** key on the keyboard when you are ready to un-pause the simulation

Flight Number	Score	Best Controlled Parameter	Worst Controlled Parameter	Parameter Most Influencing Grade	Additional Comments
1					
2					
3					
4					
5					

Additional Comments:

APPENDIX E
Instructor Experience

Instructor 1	Extensive experience in Flight simulation and flight testing procedures	
Instructor 2	5500 plus flight hours including 4000 plus instructional hours	ATP Certification
Instructor 3	7000 plus flight hours including 500 plus instructional hours	ATP Certification

REFERENCES

- ¹ FAA Advisory Circular No. 61-126 "QUALIFICATION AND APPROVAL OF PERSONAL COMPUTER-BASED AVIATION TRAINING DEVICES" Par. 5a
- ² Ludwig, J., S. Ramachandran. and W. Howse. 2002. 'Developing An Adaptive Intelligent Flight Trainer,' IITSEC-02
- ³ Mulgand, S. S., M. Asdigha, G. L. Zacharias, K. Krishnakumar and J. A. Dohme. 1995. An intelligent tutoring system for simulator-based helicopter flight training. In *Proceedings of the 1995 Flight Simulation Technologies Conference*.
- ⁴ Hoey, R.G. 1976. *Time Compression as a Means for Improving the Value of Training Simulators*. In P. Crane, D. Guckenberger, B. Schreiber, and R. Robbins (1997, August), *Above Real-Time Training (ARTT) Applied to Air Combat Skills: Final Report*, 41-65 (Tech. Report AFRL-TR-1997-0104). Mesa, AZ: Air Force Research Laboratory.
- ⁵ Kolf, J. 1973. *Documentation of a Simulator Study of an Altered Time Base*, In P.
- ⁶ Crane, P. and D. Guckenberger. 2000. *Aircrew Training and Assessment*, Ed. H.F. O'Neil and D.H. Andrews, Lawrence Erlbaum (Publisher), Mahwah, NJ, p. 153 – 164.
- ⁷ Schneider, W. 1989. Getting Smarter Quicker: Training more skills in less time, *Science and Public Policy Seminar*, Washington, DC: Federation of Behavioral, Psychological and Cognitive Sciences.
- ⁸ Barrows, A.K. and J. D. Powell. 1999. Tunnel-In-The-Sky Cockpit Display for Complex Remote Sensing Flight Trajectories, Presented at the Fourth International Airborne Remote Sensing Conference and Exhibition/ 21st Canadian Symposium on Remote Sensing, Ottawa, Ontario 21-24 June.
- ⁹ Doherty, S. and C. D. Wickens. 2001. Effects of Preview Prediction, Frame of Reference and Display Gain in Tunnel-in-the-sky Displays 11th International Symposium on Aviation Psychology, Columbus, Ohio.
- ¹⁰ Barrows, A.K., K.W. Alter., P. Enge, B.W. Parkinson and J.D. Powell. 1997. Operational Experience with and Improvements to a Tunnel-in-the-Sky Display for Light Aircraft, *Presented at the ION GPS conference*, Kansas City, MO.
<http://waas.stanford.edu/~wwu/papers/gps/PDF/andyion97.pdf>

- ¹¹ Mulder, M. 2003. An Information-Centered Analysis of the Tunnel-in-the-Sky Display, Part One: Straight Tunnel Trajectories. *The International Journal of Aviation Psychology*. 13(1), 49-72.
- ¹² Mulder, M. 2003. An Information-Centered Analysis of the Tunnel-in-the-Sky Display, Part Two: Curved Tunnel Trajectories. *The International Journal of Aviation Psychology*. 13(2), 131-151.
- ¹³ Vruels, D. and R.W. Obermayer. 1985. Human-System Performance Measurement in Training Simulators. *Human Factors*. 27(3), 241-250.
- ¹⁴ Rantanen, E.M. and D.A. Talleur. 2001. Measurement of pilot performance during instrument flight using flight data recorders. *International Journal of Aviation Research and Development*, 1(2), 89-102.
- ¹⁵ Brannick, M. T., C. Prince and E.Salas. 2002. "Reliability of Instructor Evaluations of Crew Performance: Good News and Not So Good News." *International Journal of Aviation Psychology* 14, no. 3, : 241-261.
- ¹⁶ Heath, B.E. and Crier, T. 2003. "*Comparison of Different Methods of Grading a Level Turn Task on a Flight Simulator*," Huntsville, AL.
- ¹⁷ Williams, M., 2000. Above Real Time Training and Self-Instructions for Flying on a Fidelity Validated Simulator, MS Thesis, Tuskegee University, Tuskegee, Alabama.
- ¹⁸ Wideview <http://www.wideview.it/>.
- ¹⁹ FSUIPC <http://www.schiratti.com/dowson.html>.
- ²⁰ Fltrec software www.avsim.com/hangar/utills/nav/.