January 2022

Evaluating Virtual Reality And Artificial Intelligence As Solutions For Delayed Flight Progress In Aviation Pilot Training

Ryan Paul Guthridge

EVALUATING VIRTUAL REALITY AND ARTIFICIAL INTELLIGENCE AS SOLUTIONS FOR DELAYED FLIGHT PROGRESS IN AVIATION PILOT TRAINING

by

Ryan Paul Guthridge
Bachelor of Science, University of North Dakota, 2009
Master of Business Administration, University of Texas at Dallas, 2015

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

August
2022

Name: Ryan Paul Guthridge

Degree: Doctor of Philosophy

This document, submitted in partial fulfillment of the requirements for the degree from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

DocuSigned by:

Virginia Clinton-Lisell, PhD, Chair

DocuSigned by:

Woei Hung, PhD

DocuSigned by:

Steven LeMire, PhD

DocuSigned by:

Thomas Petros, PhD

DocuSigned by:

Nicholas Wilson, PhD

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

DocuSigned by:

Chris Nelson
Dean of the School of Graduate Studies

7/26/2022

Date

PERMISSION

Title	Evaluating Virtual Reality and Artificial Intelligence as Solutions for Delayed Flight Progress in Aviation Pilot Training
Department	Educational Foundations and Research
Degree	Doctor of Philosophy

In presenting this thesis dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Ryan Paul Guthridge
July 15, 2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

There are many who helped me on this journey. I want to take a moment to provide my appreciation for their contributions to my success.

First, I wish to thank my dissertation committee. Without their guidance, I would not have been able to present the material that follows. Dr. Hung, Dr. LeMire, Dr. Petros, and Dr. Wilson served as my wise committee members. Dr. Clinton-Lisell went above and beyond as my committee chair to help me reach my goal. They all mentioned it was an ambitious field of work. They were right. It was.

To my Department Chairs, Dr. Venhuizen and Dr. Higgins, and Associate Chair, Dr. Daku. You encouraged and supported this journey, along with providing accommodations to my work balance. This was pivotal to my success and could not have been possible without your support.

Finally, to my wife, Andrea, and my two boys, Hunter and James. Your love and understanding helped me through the difficult times. Without you believing in me, I never would have made it. Pour some boujee sweet tea, open a big fig bar, and turn on some Italian cooking music. It is time to celebrate; you earned this degree right along with me!

ABSTRACT

Three studies in this dissertation examined a topic centered around delayed flight progress in aviation pilot training. Study one explored the impact of nonconcurrent flight laboratory training on the academic outcomes of collegiate aviation students, while studies two and three explored virtual reality and artificial intelligence as potential solutions to help alleviate the strain of delayed flight progress on the flight training organization. In the first study (n = 144), it was found that concurrent enrollment in an aviation classroom ground course and flight training laboratory positively impacts the mean academic block exam scores of students. In study two (n = 120), virtual reality was shown to be an effective training technology in the quantitative measure of pilot performance, as well as the qualitative measures of acceptance and adoption of the technology. Finally, the third study (n = 37) showed that an artificial intelligence-based flight instructor performs comparably to a human flight instructor, when transferring a student pilot's skills from the simulator to the aircraft. Findings from each of these studies are valuable for flight training organizations looking to find ways of better preparing their student pilots and supplementing the strain of reduced flight instructor staffing within the organization.

PROGRAM OF RESEARCH

The dissertation contained within addresses a central program of research evaluating the impact of nonconcurrent flight laboratory and academic ground courses, evaluating the efficacy of virtual reality simulation technology, and quantifying the transfer effectiveness of artificial intelligence guidance in flight simulator curricula. The common aim of this collection of three independent studies is to gain an understanding of how flight training organizations can address delayed flight training with the use of advanced training technologies. These studies have separate yet interrelated purposes, research questions, methodologies, and results that are combined into this dissertation.

In aviation pilot training, many variables make it necessary to provide flexibility for students when completing their course of training. These variables can include challenges such as, weather, financial shortcomings, academic struggles, and global crises, as we have seen with the COVID-19 pandemic. The variables inevitably slow the student's flight training in the aircraft, but often do not impede their academic progress, as they can often independently control their success in academic coursework. Because of this, students can quickly find themselves in a condition where their flight laboratory course is being completed nonconcurrent to their ground school course. One concern is that students in a nonconcurrent course of study will suffer academically, which is the first research study addressed in this dissertation.

After exploring the impact of nonconcurrent flight laboratory enrollment on academic ground course outcomes, the dissertation will explore various methods that promote pilot performance, in an attempt to provide options for students wishing to accelerate their flight laboratory course. The second article in this dissertation will evaluate the efficacy of virtual reality training devices for pilot training. Previous research has confirmed the efficacy of personal computer-based (PC-based) training technologies, which has enabled low-cost solutions for improving pilot performance. Virtual reality presents an opportunity to improve the pilot's simulated experience, while providing skills that will positively transfer to the aircraft. The second study will evaluate the efficacy of virtual reality devices and attempt to validate their use in the flight training curriculum.

Finally, the third study in this dissertation seeks to quantify the transfer effectiveness of an artificial intelligence-guided simulator pre-training curriculum for student pilots. The research will explore one possible solution of improving the rate of on-time performance by introducing a simulator curriculum to teach fundamental skills prior to beginning the flight lab. This study will use a virtual reality flight simulator, combined with an artificial intelligence-based flight instructor to guide student pilots and provide effective critique on their progress. The student pilots will progress through a self-paced curriculum in the semester immediately preceding their flight training, in an attempt to gain fundamental skills that will transfer to the flight training course.

The following pages of this dissertation contain three articles that confirm the issue of nonconcurrent flight and academic progress, then explore two options for reducing the negative impact of nonconcurrent training. These articles apply to the overarching theme of exploring how virtual reality and artificial intelligence technologies impact aviation pilot training.

Study 1

(As Prepared for Publication)

EVALUATING THE IMPACT OF NONCONCURRENT FLIGHT LABORATORY AND GROUND COURSE PROGRESS ON THE ACADEMIC OUTCOMES OF COLLEGIATE AVIATION STUDENTS

Ryan P. Guthridge

University of North Dakota, Grand Forks

Abstract

Flight training is often conducted as a two-part model, where a student completes an academic ground course to learn the knowledge and also enrolls in a flight laboratory course to apply the knowledge and skills required to earn a new certificate or rating. Often, these two parts are offered as separate courses to provide flexibility to students in the training environment. The intent is that the ground course and flight laboratory are conducted concurrently, so the students apply knowledge from the ground course during their flight training. However, external factors may delay the flight training progress in the laboratory environment, causing the student to disconnect their flight training and ground course into a nonconcurrent status. This study aims to assess the impact of concurrent versus nonconcurrent flight lab enrollment on the academic outcomes of collegiate aviation students in the classroom. The study will determine whether a student conducting flight training in their current course of study (concurrent training) performs significantly better academically than a student conducting training in a previous flight lab to their current course of study (nonconcurrent training). Quantitative data was collected in the form of academic scores on classroom block exams to evaluate the impact of students in concurrent versus nonconcurrent training environments. A series of independent sample *t*-tests were used to find consistent evidence that students in a concurrent flight laboratory perform better on block exams in their academic ground course than students enrolled in a nonconcurrent flight laboratory. The results of this research will be used to inform both educational practices within flight training departments and will assist in providing clarity to external parties interested in evaluating the impact of students completing a lab course that is nonconcurrent to their current ground course of study.

Keywords: academic outcome, concurrent, enrollment, flight, laboratory, nonconcurrent

Literature Review

Decades of research have been published concerning improving student performance, learning, and attitudes of college-level introductory science courses (Matz et.al., 2012), however little study has been done on the impact of nonconcurrent flight lab training in the aviation industry. Aviation is rooted in an educational model of providing flight lessons in a laboratory-style environment, while concurrently providing a classroom-based curriculum to learn knowledge and theory-based topics related to aviation. At the collegiate level, flight laboratories and the corresponding classroom ground courses are offered as separate components, to provide flexibility in the training environment. In some schools, students are required to concurrently enroll in the flight lab and the corresponding classroom course. However, in other schools, students are allowed to progress more rapidly through the classroom courses and may lag behind in the flight labs. This is due to multiple external factors that can delay the flight training progress in the laboratory environment. These factors can include adverse weather, flight instructor availability, or aircraft availability, to name a few.

There are a number of ways to improve student success in the flight training environment. The Airline Owner’s and Pilot’s Association (AOPA) published an article in 2015 that highlights nine habits of successful students. Many of the habits are controlled completely by the student, such as coming ready to fly, setting goals, and communication. However, there are uncontrollable factors that the AOPA study highlights, such as the ability to fly often (Deener, 2015). At the time of this publication, flight instructors are being hired to airline jobs at record rates. This leaves a shortage of qualified instructors at flight schools available to teach an increasing number of student pilots. Because of this dynamic, student progress is often dictated by their flight instructor's availability. If their availability decreases, students must find a way to become more efficient during their lessons just to remain on a reasonable timeline. Otherwise, their flight progress slows down, their flight laboratory becomes delayed, and they find themselves finishing the academic ground course without being finished with the flight laboratory course.

In 2017, advancing research in the field attempted to predict factors that attributed to student pilot success in the Part 141 collegiate flight training environment (McFarland, 2017). This research assessed the academic, cognitive, and performance attributes of 242 student pilots in a collegiate flight training program to determine which factors predicted training success. A logistic regression method was employed, which found that it was possible to predict student completion of the multi-engine flight course 73.2% of the time. The study also found a number of significant correlations amongst performance variables which indicated that academic performance is a driver of flight training success. One aspect this research assumes is that flight training and academic performance are linked in the same general timeframe. A challenge with this assumption is that many flight training schools will disconnect the flight training with the academic ground course in order to continue the student's academic progress. While the organization tracks academic progress as a key indicator of success, the student's flight training progress suffers, as they can only progress at the rate by which the flight instructor and external environment can support.

Research that expands upon existing study in the field of concurrent enrollment in lecture and laboratory comes at an optimum time with unique dynamics in the aviation industry. Current practices encourage the disconnect between laboratory and classroom instruction, such as the increased hiring of flight instructors causing a reduced ability of student pilots to maintain consistent flight training progression. In a 2016 study conducted by Lutte and Lovelace on the Regional Airline pilot shortage, the authors note that one prominent airline had a hiring target of 50 pilots for the first quarter of the year, but they only hired 28 pilots due to an acute shortage of qualified, appropriate pilots on the market. Additionally, earlier that year, this same airline was forced to cancel a scheduled training class due to a lack of qualified candidates (Lutte and Lovelace, 2016). This highlights the trend in the aviation industry, where the airlines are hiring qualified flight instructors faster than the civilian and military sectors can produce newly-qualified pilots to take their place. These dynamics influence the rate at which students complete their training. Student pilots must work one-on-one with their flight instructor to complete the flight lab lessons, whereas classroom ground courses can train upwards of 30-50 students at a time. Pressure is placed on students to accelerate the rate of their training progress, which results in students electing to continue to the next classroom ground course while they are still completing a previous flight lab course. As the student enrollment increases and flight instructor availability decreases, the chasm between flight lab progress and classroom progress increases.

Purpose of Study

The purpose of this study is to assess the impact of flight lab progress on the academic outcomes of collegiate aviation students in the classroom. It provides insight to an integral piece in assessing the impact of students not concurrently enrolled in a flight laboratory and classroom ground course. This research is a valuable addition to current research in the field that evaluates how concurrent enrollment in lecture and laboratory enhances student performance and retention. Additionally, this research helps inform the current educational methodology and training structure to help improve student academic performance in the flight training environment.

When the study was designed in 2019, airline hiring had been at an all-time high (Bureau of Transportation Statistics; BTS, 2022). Due to the COVID-19 pandemic, airline hiring was halted which resulted in a lack of pilot jobs in the industry. In turn, this resulted in a temporary surplus of flight instructors at flight schools worldwide. While this dynamic helped student pilot progress in flight schools, it is expected that flight instructors will again be rehired at airlines at greater rates than before the COVID-19 pandemic. In fact, the Bureau of Transportation Statistics shows a 2.8% month-over-month increase in airline employee hiring as of June 2022, with total employment approaching pre-pandemic levels of December 2019 (Figure 1) (Bureau of Transportation Statistics, 2022). With this expected increase in airline hiring, student pilot progress will again slow to a point where completion rates suffer in the collegiate flight training environment. Flight schools must be prepared for this effect and rely on research in the field of student success to best prepare for the capacity impact within their organization.

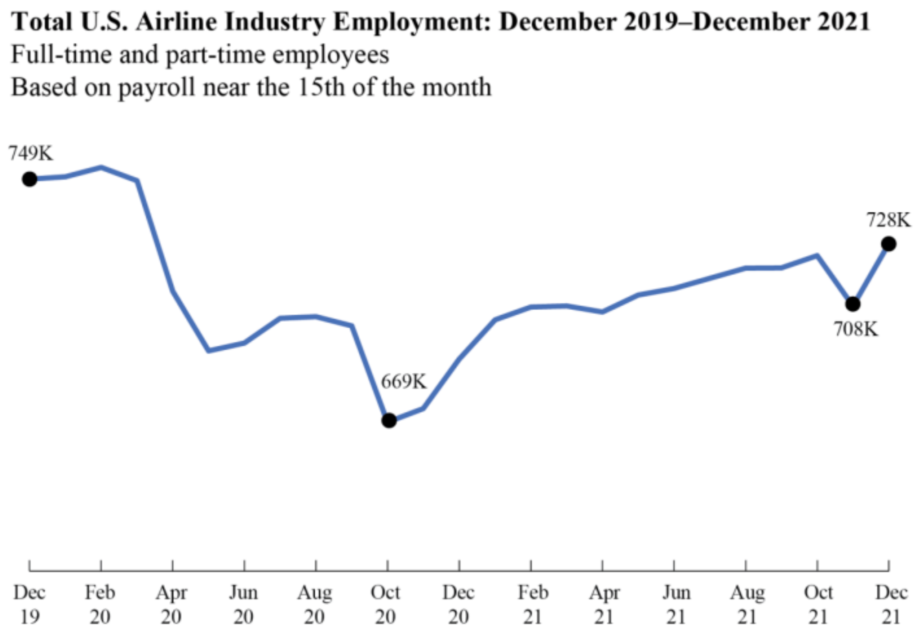

Figure 1. Total U.S. Airline Industry Employment: December 2019-December 2021 (Bureau of Transportation Statistics, 2022).

At the time of this publication, increased numbers of students enroll in flight training to fill an industry-wide pilot shortage, while facing reduced numbers of certified flight instructors available to perform their training. As student enrollment increases and flight instructor availability decreases, the chasm between flight lab progress and classroom progress is expected to widen. The results of this study will help inform existing research in the field of aviation education and include recommendations for flight training departments that are considering a nonconcurrent training model between flight lab courses and classroom ground courses.

Methods

The primary outcome of this research is to assess the academic impact of nonconcurrent flight lab courses on the academic outcomes of classroom training. A quantitative approach was used to assess the student's academic outcomes in classroom ground courses, based on their progress in the associated flight laboratory course.

Participants and Group Membership

The participants in this study were selected from students enrolled in an introductory instrument course and a flight instructor course at a midwestern university in the United States. Students were selected from these two courses to collect a dataset that was broadly representative of the total student population, as the courses are spaced at median points across the curriculum. To collect a sample from the population, data was collected from five total classes during the Fall 2020 academic semester. Within the introductory instrument course population, seven total classes were offered, which enrolled a total of 217 students. Three classes were selected from this offering, which equaled a sample size of 78 of the total 217 students enrolled during the semester. Within the flight instructor course population, four total classes were offered, which enrolled a total of 135 students. Two classes were selected from this offering, which equaled a sample size of 66 of the total 135 students enrolled during the semester.

All participants in this study successfully completed their classroom ground courses, with varying levels of progress in their flight laboratory course. Demographics of the participants can be found Table 1, which represents the combined sample population, along with the sample populations for each of the concurrent and nonconcurrent groups at the beginning of the academic semester.

At the beginning of the semester, students were assigned to groups based on their flight laboratory course enrollment. Students who were in the same flight laboratory as their ground course of training were assigned to the concurrent group, whereas students who were competing a previous flight laboratory course were assigned to the nonconcurrent group. During the semester, students were expected to continue their training in the flight laboratory course, regardless if they were completing the concurrent laboratory or the nonconcurrent laboratory. Because some students would finish the nonconcurrent laboratory between the academic block exams, their group membership would change from nonconcurrent to concurrent. Because of this factor, each block exam was analyzed independently due to the differences in group numbers at each exam. Additionally, the study accounted for block exams one through four due to the University's established last day to drop, after which many of the students in nonconcurrent laboratories dropped the academic ground course due to their delayed progress.

Table 1

Demographic Characteristics

	Combined Dataset	Concurrent	Nonconcurrent
	n = 144	n = 69	n = 75
Gender			
Male, *n (%)*	125 (86.8)	62 (89.9)	63 (84.0)
Female, *n(%)*	19 (13.2)	7 (10.1)	12 (16.0)
Academic Year			
Senior, *n (%)*	51 (35.4)	21 (30.4)	30 (40.0)
Junior, *n (%)*	49 (34.0)	24 (34.8)	25 (33.3)
Sophomore, *n (%)*	41 (28.5)	22 (31.9)	19 (25.3)
Freshman, *n (%)*	3 (2.1)	2 (2.9)	1 (1.4)
Program of Study			
Commercial Aviation, *n (%)*	121 (84.0)	60 (87.0)	61 (81.3)
Commercial Aviation & UAS Operations, *n (%)*	11 (7.6)	4 (5.8)	7 (9.3)
UAS Operations, *n (%)*	9 (6.3)	4 (5.8)	5 (6.7)
Commercial Aviation & Management, *n (%)*	3 (2.1)	1 (1.4)	2 (2.7)

*Note. D*emographics were collected at the beginning of the academic semester.

Quantitative Study

The purpose of the quantitative study was to determine the degree at which nonconcurrent flight lab training impacts the academic outcomes of students in the classroom ground course. Academic performance data was collected in the form of block exam scores. The structure of the academic ground courses was to provide block exams that are comprehensive to a building block of learning in that course. The block exams were spaced at approximately one-month intervals during the Fall 2020 academic semester. Because of this, each of the two courses were evaluated separately during the data analysis phase, due to the difference in evaluation content and criteria for each of the respective block exams. The block exam scores were aggregated into populations based on concurrent and nonconcurrent flight lab enrollment at the time the participant took the Block Exam.

A series of independent samples *t*-tests were conducted to evaluate the mean difference between students enrolled in a concurrent flight laboratory and a nonconcurrent flight laboratory. Eight *t*-tests were conducted in total, which compared each of the four block exams for two separate academic ground courses during the Fall 2020 semester.

Results

The Introductory Instrument Course

The introductory instrument course is offered immediately after the student finishes their Private Pilot training. In this course, a total of 217 students enrolled during the Fall 2020 semester. This study sampled three classes of the total population of the introductory instrument course, which equaled 78 students (35.9%) of the total population. In this sample, 41 students (52.6%) began the flight laboratory concurrently with the academic ground course. The remaining 37 students (47.4%) were still finishing the Private Pilot flight laboratory, and were considered to be in a nonconcurrent laboratory.

Students in this academic course spend Block One reviewing content related to the Private Pilot course, which typically garners higher results during the Block One exam since the students have recently trained on this content to proficiency prior to enrolling in the introductory instrument course. Subsequently, the course proceeds to cover topics of flight instrument systems, methods of basic attitude instrument flying, and navigation systems. Blocks Two through Four offer a more in-depth study of topic areas and may be considered “new content” for the purposes of learning the material. Because of this, the results of Block Exams Two through Four could be related to a traditional academic course that offers new content for all blocks of learning.

In this study, there was no significant effect for Block One exam scores, $t(76) = 1.191$, $p = .237$, despite students in a concurrent lab ($M = 88.41$, $SD = 8.11$) scoring higher than students in a nonconcurrent lab ($M = 86.22$, $SD = 8.17$). For Block Two exam scores, students in a concurrent lab ($M = 88.94$, $SD = 9.15$) scored significantly better than students in a nonconcurrent lab ($M = 80.07$, $SD = 9.59$), $t(76) = 4.065$, $p = .001$. For Block Three, students in a concurrent lab ($M = 89.38$, $SD = 7.56$) scored significantly better than students in a nonconcurrent lab (*M* = 78.44, *SD* = 20.01), *t*(76) = 3.517, *p* = .001. Finally, for Block Four exam scores, students in a concurrent lab (*M* = 80.76, *SD* = 10.11) scored significantly better than students in a nonconcurrent lab (*M* = 75.25, *SD* = 11.66), *t*(76) = 2.020, *p* = .047.

In the results above, the Block One exam presumably did not show significance due to the nature of the content of the Block One exam. Content on this exam is a review of material that was recently completed by the students in the course immediately preceding this course. For the remainder of the Block Exams, significance was found between the concurrent and nonconcurrent groups. Figure 2 and Table 2 show the results of each block exam score for the introductory instrument course.

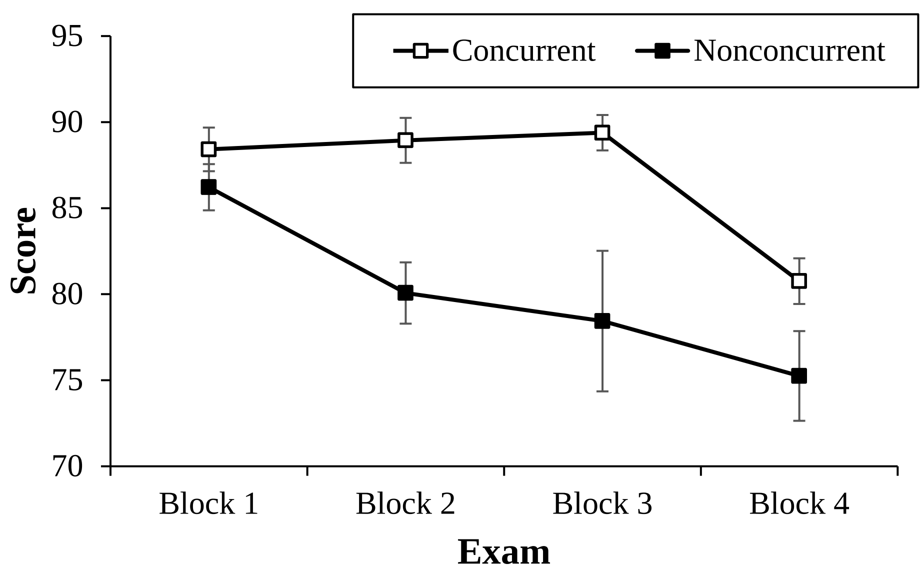

Figure 2. Introductory Instrument Course Block Exam Scores

Table 2

Introductory Instrument Course Block Exam Scores

	Concurrent Lab *(n)*	Nonconcurrent Lab *(n)*	*p*
Block One, *score (n)*	88.41 (41)	86.22 (37)	.237
Block Two, *score (n)*	88.94 (49)	80.07 (29)	.001*
Block Three, *score (n)*	89.38 (54)	78.44 (24)	.001*
Block Four, *score (n)*	80.76 (58)	75.25 (20)	.047*

Note. * $p < .05$

The Flight Instructor Course

The flight instructor course is offered immediately after students finish a course in commercial multi-engine flying. Students that enroll in a concurrent flight laboratory learn how to teach fundamentals of aviation instruction in a single-engine aircraft, while students in a nonconcurrent laboratory course are learning how to master the pilot-in-command responsibilities of a multi-engine aircraft. These courses are significantly different in structure and content, which likely explains the consistent difference in scores on each block exam.

The initial split of students in nonconcurrent and concurrent flight laboratories was wider in this course, largely due to the complex nature of the preceding multi-engine course. The multi-engine course requires uniquely qualified flight instructors, which slowed down the progress of the population of students planning to enroll in the flight instructor academic ground course. In this course, a total of 135 students enrolled during the Fall 2020 semester. This study sampled two classes of the total population of the flight instructor course, which equaled 66 students (48.9%) from the total population. In this sample, 28 students (42.4%) began the flight laboratory concurrently with the academic ground course. The remaining 38 students (57.6%) were still finishing the multi-engine flight laboratory and were considered to be in a nonconcurrent laboratory.

Students in the academic course will spend time learning fundamentals of instruction, which includes topics related to lesson planning, content delivery, student evaluation, and assessment. These topics are combined with technical subject areas related to general flight, including aerodynamics, aircraft performance, systems, flight planning, and flight maneuvers. Generally, these topic areas have been previously learned by the students, however they are now expected to learn and teach these topics at an instructor's level of knowledge. For the purposes of this course, all blocks of learning could be considered "new content" from a fundamentals of instruction perspective, even though there are a number of content areas that are familiar to students, in the form of technical subject areas they have previously learned.

In this study, all Block Exam scores showed significance, with similar raw score differences between the concurrent and nonconcurrent groups on each Block Exam. For Block One exam scores, students in a concurrent lab (M = 89.46, SD = 5.75) scored significantly better than students in a nonconcurrent lab (M = 85.17, SD = 8.06), $t(64)$ = 2.402, p = .019. For Block Two exam scores, students in a concurrent lab (M = 90.65, SD = 5.39) scored significantly better than students in a nonconcurrent lab (M = 86.86, SD =

7.90), $t(64) = 2.244$, $p = .028$. For Block Three exam scores, students in a concurrent lab ($M = 89.87$, $SD = 4.53$) scored significantly better than students in a nonconcurrent lab ($M = 84.36$, $SD = 6.12$), $t(64) = 4.208$, $p = .001$. Finally, for Block Four exam scores, students in a concurrent lab ($M = 87.37$, $SD = 5.99$) scored significantly better than students in a nonconcurrent lab ($M = 84.36$, $SD = 5.61$), $t(64) = 2.023$, $p = .047$. Figure 3 and Table 3 show the results of each block exam score for the flight instructor course.

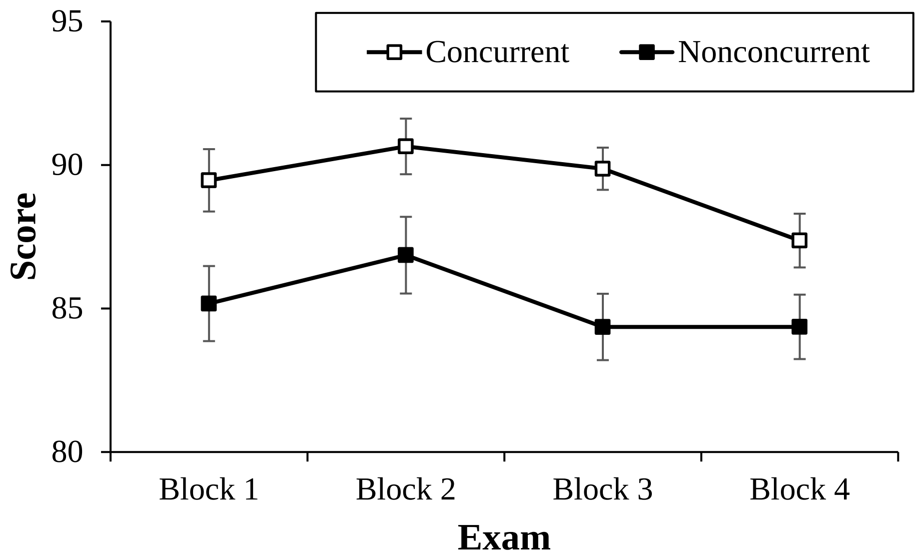

Figure 3. Flight Instructor Course Block Exam Scores

Table 3

Flight Instructor Course Block Exam Scores

	Concurrent Lab *(n)*	Nonconcurrent Lab *(n)*	*p*
Block One, *score (n)*	89.46 (28)	85.17 (38)	.019*
Block Two, *score (n)*	90.65 (31)	86.86 (35)	.028*
Block Three, *score (n)*	89.87 (38)	84.36 (28)	.001*
Block Four, *score (n)*	87.37 (41)	84.36 (25)	.047*

Note. * $p < .05$

Discussion

The key finding of this study is that concurrent enrollment in aviation ground course and flight training laboratory positively impacts academic outcomes. As the Aviation industry climbs out of the COVID-19 pandemic and hires airline employees at pre-pandemic rates (Bureau of Transportation Statistics, 2022), these findings provide important guidance to flight training organizations on methods that hinder student pilot academic success. These findings are particularly important when considering methods to alleviate organizational capacity demands when faced with a flight instructor shortage. Additionally, as incoming student enrollments increase, these findings provide guidance to evaluate alternative methods to providing an appropriate training structure that ensures the academic success of students enrolled at the flight school.

One consideration this study addresses, is the range of courses and experience offered by a flight training organization. When pursuing a career as a professional pilot, each flight training course provides a different level of intensity due to the wide range of knowledge and skills required across the curriculum. While looking at the programmatic requirements of the flight training curriculum, one might consider the initial private pilot course and the flight instructor course as the most intensive training courses offered. Alternatively, the introductory instrument course might be considered one of the courses with the least training intensity. In any case, the findings of this study highlight the importance of maintaining concurrent enrollment in a flight laboratory that matches the academic ground course.

Nearly all block exams showed statistical significance, with the one exception being the Block One exam in the introductory instrument course. As stated previously, this exam is a review of material previously learned by students in the course immediately preceding the introductory instrument course. Because of this, it was expected that all students would perform similarly on the Block One exam, regardless of concurrent or nonconcurrent laboratory status.

When considering the raw score differences amongst all block exams in the data set, students in a concurrent flight laboratory scored 5.5% higher on block exams than students in a nonconcurrent flight laboratory, on average. Functionally, this would be equivalent to a student receiving a grade of A in the class, versus a student receiving a B+. Alternatively, this could be the difference between a student successfully passing the academic ground course and a student being required to retake the same course due to a failing grade.

The findings of this study show the importance of maintaining concurrency between a student pilot's flight laboratory and the associated academic ground course. Research has shown that students who engage in well-designed laboratory experiences develop problem-solving and critical-thinking skills, as well as gain exposure to reactions, materials, and equipment in a lab setting (ACS, 2022). However, it is important that students apply the knowledge in a timely manner, which is the primary reason why a student enrolled in a nonconcurrent laboratory suffers academically. These students are applying knowledge from a previous academic course in their laboratory, while attempting to learn new content in their current academic ground course. This disconnect may be detrimental to a student's academic success, and therefore every effort should be made to avoid nonconcurrent laboratories during their flight training.

Limitations

Limitations of this study center around the dynamics related to group membership and the reasons for switching from a nonconcurrent to a concurrent laboratory status. There are many reasons that a student becomes delayed in their flight training. Natural causes may include weather, flight instructor availability, or aircraft availability, to name a few. Other variables may be more undetectable, including stress, fatigue, financial hardship, or relationship struggles. It is important to note that these potentially confounding variables were outside of the scope of this research and not accounted for in the dataset.

Finally, when a student finds themselves in a nonconcurrent laboratory status, they may take on an alternative approach to their academic success, versus students in a concurrent laboratory. For instance, some students in a nonconcurrent laboratory may put more effort into remaining proficient in the knowledge and skills required by the previous academic course, in order to ensure their success in the nonconcurrent laboratory lessons. These students may suffer academically in the concurrent course, since they are choosing to focus on different content. Alternatively, students in a nonconcurrent laboratory may choose to focus more intensely on the new content of the concurrent course, in order to not fall behind and suffer in the classroom. Academic motivation was not collected during this study and was not accounted for during the analysis.

Implication for Practice

The results of this study show that value should be placed in maintaining a concurrent flight laboratory and classroom ground course with all students in the curriculum. Additionally, this research shows that students may suffer academically if they accelerate their classroom ground courses without first completing any previous flight laboratory courses that are required by the curriculum. Risks to an educational model that provides nonconcurrent flight laboratory and classroom ground training are a significant decrease in classroom academic performance.

Study and research of this topic in the aviation industry is integral to maintain and bolster the pilot pipeline, while maintaining the proficiency and knowledge standards employed by the industry. Beyond the research presented in this paper, it is suggested to employ these statistical methods on aviation training models outside of the primary flight training environment. These could include recurrent training and initial type rating training. Additionally, researchers may wish to include academic motivation as an additional variable when choosing to replicate this study. For instance, in a recent study by Wilson and Stupnisky (2022), the authors use the Academic Motivation Scale (AMS; Vallerand et al., 1992) to evaluate for differences in motivation between students who enrolled in either a blended course or an online, asynchronous section of a senior-level advanced aircraft systems course. A similar methodology could be employed to evaluate the differences in motivation for students in a nonconcurrent and a concurrent flight laboratory course.

Finally, research consideration should be explored in providing a structured, self-paced pre-training course for student pilots that may help accelerate and increase the proficiency of training in the flight lab courses, thus increasing the probability of maintaining concurrency between the flight lab and classroom ground courses within the flight training curriculum. Finally, future research should be conducted to evaluate the efficacy of low-cost flight simulation technologies, that could be used to support a self-paced training curriculum by student pilots, which would not be reliant on flight instructor availability for a successful outcome.

References

ACS, American Chemical Society. (2022). Importance of hands-on laboratory science. Retrieved June 3, 2022, from https://www.acs.org/content/acs/en/policy/publicpolicies/education/computersimulations.html

Bureau of Transportation Statistics. (2022). U.S. Cargo and Passenger Airlines Add 20,049 Jobs in December 2021, Reaching COVID-19 Pandemic High; Employment Remains 2.9% Below Pre-Pandemic December 2019. United States Department of Transportation. https://www.bts.gov/newsroom/us-cargo-and-passenger-airlines-add-20049-jobs-december-2021-reaching-covid-19-pandemic

Deener, S. (2015). Nine Habits of Successful Students. Airline Owner's and Pilot's Association. https://www.aopa.org/news-and-media/all-news/2015/january/13/habits-of-successful-students

Lutte, R.K., & Lovelace, K. (2016). Airline Pilot Supply in the US: Factors Influencing the Collegiate Pilot Pipeline.

Matz, R. L., Rothman, E. D., Krajcik, J. S., & Banaszak Holl, M. M. (2012). Concurrent Enrollment in Lecture and Laboratory Enhances Student Performance and Retention, Journal of Research in Science Teaching, 2012-May. Https://Eric.Ed.Gov/?Id=EJ989475. https://eric.ed.gov/?id=EJ989475

McFarland, M.R. (2017). Student Pilot Aptitude as an Indicator of Success in a Part 141 Collegiate Flight Training Program. Kent State University College of Education.

https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=kent1492088859648498

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. Educational and Psychological Measurement, 52(4), 1003–1017. https://doi.org/10.1177/0013164492052004025

Wilson, N. & Stupnisky, R. (2022). Examining differences in aviation student motivation during blended versus online asynchronous courses. Collegiate Aviation Review International, 40(1), 95-112. Retrieved from http://ojs.library.okstate.edu/osu/index.php/CARI/article/view/8564/7783

Study 2

(As Prepared for Publication)

EVALUATING THE EFFICACY OF VIRTUAL REALITY (VR) TRAINING DEVICES FOR PILOT TRAINING

Ryan P. Guthridge

University of North Dakota, Grand Forks

Virginia Clinton-Lisell, PhD

University of North Dakota, Grand Forks

Acknowledgements

This study was funded by a grant from the John D. Odegard School of Aerospace Sciences at the University of North Dakota. The depth of study and access to costly flight training devices was greatly improved with this support.

The researchers were fortunate to have the help of two outstanding research assistants, Alexis Hesse and Caroline Kelley. Their ability to produce a professional work product is unmatched and we were grateful to have them as part of our research team.

Abstract

Virtual Reality (VR) technology is a quickly advancing field that has many documented benefits, including highly detailed environments, accuracy to the real world, and low cost of entry in the flight simulation market (Radianti et. al., 2019). At the time of this study, VR technology has not been well tested or widely accepted in the aviation industry. This research project seeks to evaluate the efficacy of virtual reality (VR) training devices for use in a pilot training program. The hypothesis is that pilots who train in virtual reality simulators will perform significantly better than pilots who train on PC-based flight simulators and thus will transfer these skills more efficiently to the actual aircraft during flight training. This specific study will be conducted on beginning-level instrument pilots while performing a visual traffic pattern at an airport. Quantitative and qualitative data was collected to support the research, conducted on students in flight simulators. A one-way ANOVA was used to evaluate the equivalence of each group in the study based on previous flight and VR experience. Then, a one-way ANOVA was conducted on pre-test/post-test gain scores to compare each training group, as well as a post hoc Tukey HSD to conduct multiple comparisons and evaluate mean differences between the groups. The results show that participants who train in a VR simulator perform similarly to students who conduct training in a PC-based simulator. Additionally, both training groups performed significantly better than the control group, which conducted no training between the pre-test and post-test. Finally, survey data was evaluated to find that students who trained in VR simulators felt as if they performed better on the post-test than the pre-test. Comments from the students indicated that most felt as though VR simulators could be an acceptable training technology for use in the flight training curriculum. These results will help inform flight training organizations who are considering new technology that provides a low-cost and high-value alternative to costlier, fixed-based simulators.

Keywords: flight training, flight simulators, mixed methods, virtual reality

Literature Review

Virtual Reality is a computer-generated simulation of a three-dimensional environment, through which the user can interact similarly to the real world. This technology has advanced to the point of mainstream use in our daily lives through our phones, tablets, and computers, but the new technology is not without its challenges. One challenge of this new technology is the level of fidelity it provides to the user, since a smooth and accurate visual environment is imperative for flight training accuracy. Previous research to evaluate the fidelity of flight simulators has centered around three main themes, which include how the simulator replicates the real-world environment (Bradley & Abelson, 1995), the simulator's visual field of view (Reweti, Gilbey, & Jeffrey, 2017), and how the simulator replicates the sensations of flight (including motion and tactile feedback) (Duncker, 1938).

Replicating the Real-World Environment

One feature of a flight simulator that makes a substantial impact on a pilot's skill transfer to the aircraft is the simulator's ability to replicate the real-world environment. The FAA has also recognized the importance of this feature by establishing a requirement that flight simulator "control inputs should be reflected by the flight instruments in real time and without a perceived delay in action." (FAA, 2018). This particular feature in flight simulators has been explored in detail, with guidance provided that establishes a threshold for simulator capabilities that impact a pilot's ability to establish consistent control input in the simulator.

In a study published in 1995, Bradley and Abelson review some of the factors that determine how well a simulator captures the actual experience of flight. One of the main concerns they address is the issue of computer frame rates (the rate in which a computer provides a new image of the real-world environment) versus a pilot's ability to accurately control the aircraft. They indicate that "to the extent that a flight simulator is not entirely realistic, it must be due to one of two things: hardware limitations or that the underlying theory used in the program is in some respect incomplete or incorrect." This is particularly important, because the student must "master the intricate feedback relationship between his or her control inputs and the resulting changes in the outside visual environment and the instruments." While simulator training in the IFR environment is more difficult to master, it is simulated flight in VFR conditions that is more difficult to program. In a simulated VFR environment, there is a requirement to "generate and display a constantly changing out-the-window view (i.e., to do real-time animation), which taxes the computational capability of the computer" (Bradley & Abelson, 1995). This research explores the technological requirements of a simulator in order to reduce a delay in frame rates and provide the pilot with a seamless visual experience in the simulator.

When discussing the issue of frame rates in flight simulators, Bradley and Abelson state that in order to produce the impression of an aircraft moving through space, the program must create and display frames at a rate of 15-30 frames per second. Unfortunately, this is not always possible due to the limitations in the processing speed of the computer system and the demands made on the simulation at a particular point in time. For instance, in programs that have highly-detailed scenery, the frame rate may drop as low as 4-6 frames per second during computationally-intensive periods. This means that over 75% of the information in the simulation is being omitted. Where this comes into play is at low altitudes, close to the runway surface, where the simulated environment is highly-detailed and changes rapidly. This is particularly detrimental when attempting to control an aircraft accurately, since pilots must make precise movements, observe the visual effects of those movements, and adjust them as necessary. Now imagine introducing a delay between when a control movement is supplied and when the resulting visual feedback is displayed on the computer monitor. When the visual environment does not update fast enough, it can lead the pilot to make inaccurate corrections and either over-control or under-control the aircraft. “Since judging the effects of the correction requires judging rates of change over time, the fewer the frames that are displayed each second, the longer it will take the pilot to properly assess the effects of the correction” (Bradley & Abelson, 1995, p. 157).

To display this effect, Bradley and Abelson conducted an experiment that simulated delayed sensory feedback. Subjects were required to trace patterns, such as the stars in Figure 1, and imposed either no delay, a 0.52 second delay with continuous visual feedback, or an intermittent visual feedback of 0.17 second (to simulate a 6 frames per second delay). What the research showed was with longer delays in visual feedback, the subjects continued to apply the input until visual feedback was received. In the flight simulation example, a pilot “expects the control input to produce a more or less immediate effect, and when it doesn’t, the natural tendency is to supply additional control input to get the aircraft moving in the desired direction” (Bradley & Abelson, 1995, p. 157). This shows that there is a tradeoff between computational performance and its ability to supply a highly detailed environment.

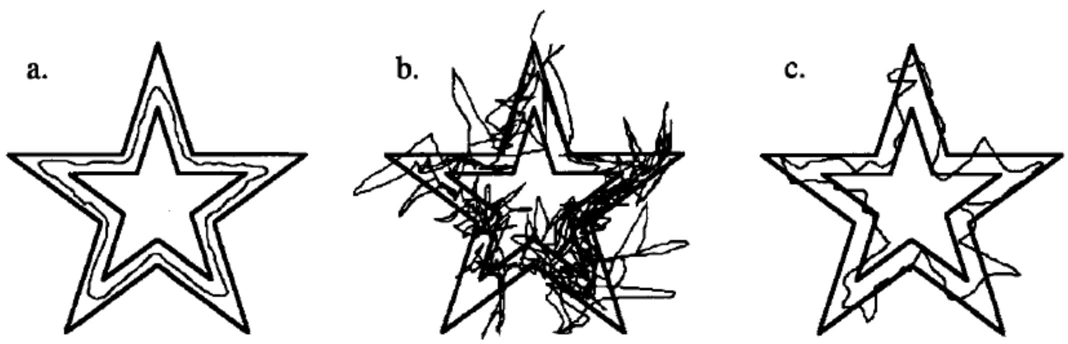

Figure 1: Star tracings under conditions of (a) no delay, (b) a 0.52 second delay with continuous visual feedback, and (c) intermittent visual feedback at 6 fps and frame durations of 0.17 second (Bradley & Abelson, 1995).

Bradley and Abelson (1995) do an excellent job of describing the challenging solution:

> To overcome the problems of delayed sensory feedback, the processing speeds of desktop systems will have to improve enough that any delay is less than a few hundredths of a second. However, this optimism is offset by the tendency of programmers to demand more than the current processor technology can support. In the future, programmers would be well advised to give more attention to the tradeoff between maintaining

> aircraft responsiveness and portraying a richly detailed visual environment (p. 158).

Virtual reality technology faces a similar challenge. As the new technology provides highly-detailed, 360 degree environments, software developers have an opportunity to increase the visual detail, which is computationally expensive and could result in delayed frame rates. This is similar to the problems seen by Bradley and Abelson in 1995 with PC-based technology.

Visual Field of View

One of the first noticeable limitations of a flight simulator is the limited vision created by the computer monitor or projector screens. Typically, PC-based flight simulators provide a visual environment immediately in front of the pilot, with a view angle of approximately 70 degrees. Additionally, since the bottom portion of the monitor is used to display the instrument panel of the aircraft, the vertical visual field is also substantially reduced. It is important to note that the horizontal span for unrestricted binocular vision is a visual field extending 200 degrees, which is a common component of more advanced (and costly) fixed-based flight simulators. Because of the limited visual field of view, even experienced pilots find it difficult to fly accurately using a PC-based simulator, and overshooting and undershooting are common when turning. Research by Reweti, Gilbey, and Jeffrey (2017) addresses this concern by comparing pilot performance in two groups: one that trains in a PC-based simulator and one that trains in a fixed-based simulator.

In the actual aircraft, pilots have a near-360-degree field of view, which allows them to scan for traffic, provide adequate spacing to obstacles, and accurately perceive the aircraft's proximity to the runway environment during the approach and landing phase of flight. In the environment provided by a flight simulator, the visual field of view is significantly reduced. In Reweti, Gilbey, and Jeffrey's study, two industry-accepted versions flight simulators were compared, the PC-based simulator (providing 62 degrees of view) and the fixed-based flight simulator (providing 170 degrees of view). Identical training was provided in each simulator type, with the pilots' performance being compared between a pre-test and post-test flight, which occurred on either side of the training curriculum. In this study, "no overall evidence was found that a fixed-based flight training device performed better than a PC-based simulator when used to train pilots on a VFR flight maneuver." More specifically, the researchers found no difference in the efficacy of a PC-based simulator and a fixed-based simulator. However, the use of both simulator types "demonstrated a significant improvement in VFR task performance compared to a control group that received no simulator training" (Reweti et. al., 2017).

The findings of Reweti, Gilbey, and Jeffrey (2017) are significant in demonstrating the efficacy of a low-cost PC-based simulator compared to its high-cost counterparts, as well as the efficacy of flight simulation devices compared to a control group that received no simulator training. By providing data on the efficacy of low-cost simulators, this research opens the door for a broader use of flight simulators by flight schools of all sizes, since cost is no longer a barrier to entry. Additionally, this research provided much needed data to support legislation that allows basic aviation training devices to be authorized for use in flight training. From the FAA’s Advisory Circular (FAA, 2018), they explain the general requirements of a basic aviation training device to include control systems similar to what a generic airplane would use and software that replicates generic aircraft flight dynamics. To this end, even though the basic aviation training device may not replicate the actual training aircraft or provide a highly-immersive visual field of view, it is allowed to be used to supplement the training time requirements of student pilots seeking a new certificate or rating.

Replicating the Sensations of Flight

While flight simulators take on various configurations, one aspect that has been commonly-accepted as a “must have,” is the simulator’s ability to replicate the sensation of flight. This topic addresses the concern that many flight simulators lack the ability to provide full motion or to provide force-feedback on the pilot’s control inputs. These two topics will be covered here, along with highlighting research and theories derived from Gestalt psychology as it relates to the sensations of simulated flight.

The first topic of concern is simulated motion. Historically, federal regulators around the world have recognized large, costly, full-motion simulators as a suitable solution for replicating “real-world” flight. In many cases, regulators have authorized full-motion simulators to be used as a replacement for the actual aircraft in training curriculum, to the degree that some may even be used to record takeoffs and landings for pilot currency purposes. The problem is that large, costly, full-motion simulators are owned by airlines, lessors, or third-party training organizations, and are not available to be used by the majority of the flying public. To this degree, research has been conducted to determine if the motion feature of flight simulators actually makes an impact on the performance of pilots.

Karl Duncker explores the topic of motion in his 1938 publication, titled “Induced Motion.” He states:

> In normal vision, objective motion can be experienced both when a moving stimulus traverses the resting retina and when the eye itself follows the stimulus. There is, however, a type of perceived motion quite different from all others, and this is the so-called “induced motion.” When, for example one is sitting on a railway coach and a nearby train moves, it seems for a time as if one’s own train were moving in the opposite direction. This is a case of induced motion. (Duncker, 1938, p. 163)

Induced motion plays a key role in simulating the perceptions of flight through three-dimensional space. This feature was explored by Go and colleagues (2000) in their research designed to assess the degree by which motion affected the training of skills and, most importantly, the transfer of those skills to the airplane (p. 2). The research conducted by Go and colleagues was an attempt to assure that FAA legislation for simulator training requirements promote full transfer of pilot performance between simulator and airplane, without unnecessarily driving up cost. This is an important research question, as the motion feature on flight simulators comes with a substantial increase in costs, which may be prohibitive for some flight training schools. Likewise, if it is proven that motion provides no statistically significant change in performance transfer between the simulator and the aircraft, flight training schools would be able to more feasibly offer training in low-cost aviation training devices, rather than spending unnecessary costs on a full-motion device.

The Go and colleagues (2000) study was unique, in that it compared both objective and subjective performance of pilots when using simulators with, and without, motion. The motion and non-motion groups were tasked with flying a variety of maneuvers while data was collected on their objective performance. Additionally, during the tasks, instructor evaluators were tasked with conducting subjective evaluations of the pilots' performance. Upon completion of the tasks, the objective data and subjective evaluation grades were compared to determine the effectiveness of motion on pilot task performance. What was found was that "platform motion had no effect on the grades that were provided by the instructor evaluators." Additionally, "no statistically significant differences in improvement from first to last training trial were found between groups for any of the measures conducted in the study. This suggests that the platform motion did not affect the training progress of the pilots" (Go et. al., 2000, p. 3).

The second feature to consider when discussing the sensation of flight is that of force feedback and its impact on a pilot's tactile response. To describe this sensation, Rock and Victor (1964) conducted an experiment that observed a subject's response when presented with two properties of the same object that caused conflict in the senses.

In other words, “if contradictory information is given to two senses of an observer about the properties of an object, what will be his experience?” (Rock & Victor, 1964, p. 594). This study was integral in explaining the feature of tactile feedback when flying a training device largely by visual reference to the simulated environment.

The study utilized three different experiments; (1) visual comparison only, (2) tactile comparison only, and (3) a different method in which the subject was asked to draw a picture of the same shape as the sample provided. The subjects were provided the ability to touch a three-dimensional object while simultaneously viewing that object through a transparent optical element, which compressed the image along its horizontal axis only. These experiments compared conflicting information across the tactile and visual senses, which are the two senses used when flying a simulator using the flight controls and a visual representation of the external environment. What the researchers found was that in all three experimental conditions, the visual impression was completely dominant. “In other words, vision is so powerful in relation to touch, that the very touch experience itself undergoes a change” (Rock & Victor, 1964, p. 595). The object actually feels the way it looks and this is the reason why the researchers believed that the subjects were unaware of a conflict in the visual and tactile sensations observed during the experiments.

What the research by Rock and Victor (1964) showed, was that when conflict exists between the tactile and visual senses, the subject will believe the visual sensation as the truth. So much to the point where the subject believes that what they are feeling (even when incorrect) represents the real-world condition. When applying this research to flight simulation, the senses felt through both force-feedback on the flight controls, as well as platform motion, do not present a significant improvement in the transfer of piloting skills between the simulator and the aircraft. In some cases, the tactile feedback provided through the controls or through platform motion could be considered a non-essential feature, as the visual senses play such a strong role in creating insights from the sensations received, that the subject believes that what they are feeling matches what they are visually observing.

Virtual Reality Flight Simulators

The use of flight simulators at all levels of pilot training has been proven to positively impact the performance of a student during their training course, however the fidelity of simulation and lack of immersion often limits the realism of training. With the advent of virtual reality (VR) flight training simulators, an individual can fully immerse themselves in the virtual world and the transfer of skills should be nearly identical. The information gained from conducting a robust mixed methods study would help evaluate the efficacy of VR technology for use in flight training, as compared to PC-based flight simulators that are currently being used in today's training environment. Likewise, qualitative survey responses will increase our understanding of a pilot's willingness to accept flight simulator technology as a beneficial addition to the training curriculum and identify the perception of the realism of the simulated environment to the actual real-world training environment.

Purpose of Study

This research is an integral piece in quantifying the impact of simulator-based training solutions in aviation pilot training. Particularly, with the advent of VR simulators and the gamification of training, future curriculum will be developed that sufficiently enhances the quality of training provided to the pilot, thus reducing the required amount of training with a Certified Flight Instructor (CFI) in a flight training aircraft. The reduction in required training is significant and comes at an optimum time, with a measurable shortage of flight instructors available to train new students. In turn, this will cause an increased backlog of students waiting to receive dual instruction from flight instructors in both the civilian and military sectors. In a 2016 study conducted by Lutte and Lovelace on the Regional Airline pilot shortage, they note that one prominent airline "had a hiring target of 50 pilots for the first quarter of the year, but they only hired 28 pilots due to an acute shortage of qualified, appropriate pilots on the market." Additionally, earlier that year, this same airline was "forced to cancel a scheduled training class due to a lack of qualified candidates" (Lutte and Lovelace, 2016, p. 55). This highlights the trend in the aviation industry, where the airlines are hiring qualified flight instructors faster than the civilian and military sectors can produce newly-qualified pilots to take their place.

The use of flight simulators as a principle component of flight training depends upon a few key factors to ensure their success. First, the Federal Aviation Administration (FAA) ultimately has oversight and provides regulation pertaining to what types of flight simulation devices are allowed to be used to satisfy the training time requirements for initial and recurrent training. Second, the industry must be willing to accept and adopt the flight simulation device in order for it to be used in training. Lastly, there must be sufficient oversight to ensure the flight simulator does not introduce unintended risks to the learning outcomes of students.

The first topic of FAA oversight and regulation is a benefit to using flight simulation technologies during training. While the FAA has historically been slow to adopt to new technologies, they recently funded research to determine which simulator features promote full transfer of pilot performance between simulator and airplane, without unnecessarily driving up cost. A result of this research was the publication of an Advisory Circular titled *FAA Approval of Aviation Training Devices and their Use for Training and Experience* (FAA, 2018), which describes the minimum requirements for flight simulators to be certified for use during initial and recurrent pilot training. This is a benefit for the use of technology in the Aviation industry, since the FAA recognizes the value of various types of flight simulation technologies and provides a benefit to students who use them as a part of their initial and recurrent training curriculum. Additionally, since the FAA now authorizes the use of basic aviation training devices, flight schools can realize the benefits of using flight simulators without being required to spend an excessive amount of cost on a fixed-base training device.

The second topic of industry acceptance and adoption is also a benefit to using flight simulation technologies during training. Flight simulators have been used to train pilots since the 1930's when the Link Trainer was developed for pilots learning how to fly the Grumman Avenger in World War II (McElhiney, n.d.). Since then, flight simulators have been adopted to help teach student pilots in a low-cost and low-risk environment. Because of the long-standing nature of using flight simulators for training pilots, the Aviation industry has widely accepted the flight simulator as a supplement to training in the actual aircraft. While the acceptance of flight simulators as a technology is widespread, the level by which simulators are included in training curriculum often varies by what type of flight simulator is available to the instructor. This is where the research presented above plays a significant role in the adoption of varying levels of flight simulators into the training curriculum. Flight simulators have been observed to demonstrate a significant improvement in task performance compared to groups that conduct no simulator training (Reweti, et. al., 2017). Included in this are generic PC-based flight simulators, which have been shown to provide a significant improvement in task performance, even when the simulator does not replicate the actual aircraft. By following the outcome of research, the FAA has allowed the use of basic aviation training devices in initial and recurrent pilot training, which is a benefit to flight schools looking to adopt flight simulators into their existing curriculum at a lower-cost than previously expected with fixed-based and full motion flight simulators.

To address these topics and to evaluate the efficacy of VR technology for pilot training, the following research questions were developed:

1. How do virtual reality flight simulators compare to PC-based flight simulators, in respect to pilot performance?
2. How do students feel the virtual reality compares to a traditional flight training device?
3. Would the students accept virtual reality as a suitable alternative to traditional flight training devices?

Methods

The primary outcome of this research project is to evaluate the efficacy of VR training devices for pilot training. A mixed methods approach was used to quantitatively assess the pilot's task performance after training in a VR training device and qualitatively evaluate the pilot's perception of the VR technology as it relates to acceptance and adoption in the flight training environment. VR simulator technology was compared to PC-based simulator technology and these types were compared with a control group that conducted no training. Both methods of research are described in detail below.

Participants

The participants in this study were selected from beginning-level instrument pilots at a collegiate aviation university in the United States. In order to participate in this research, students were required to have successfully passed their Private Pilot check ride. Participants were excluded from the research if they had not yet obtained their Private Pilot certificate or if they did not meet the prerequisite to enroll in the beginning-level instrument course. The prerequisite to enroll was the successful completion of the

Private Pilot academic ground course. This maintains the validity of the pool of applicants to comparable levels of measured flight proficiency, as determined by the FAA Airman Certification Standards, which governs the standards for earning pilot licenses in the United States.

For two consecutive academic semesters (Fall 2021 and Spring 2022), three classrooms were selected and assigned to training groups. A quasi-experimental design was applied, which allowed one classroom to serve as the control (no training) group, one classroom conducted PC-based training, and one classroom conducted VR training. Students were instructed to only fly their training type (VR, PC, or none). By applying quasi-experimental methodology, the simulators could be set to an identical configuration and students would not have the awareness to train using a different type of simulator configuration, due to all of their peers training with the same simulator configuration.

Experimental Procedure

During the first two weeks of the academic semester, students completed the experimental procedure for this research study. During the first week, students received their group assignment and completed a pre-test flight in a Frasca Aviation Training Device (ATD). During the second week, students in a training group conducted three practice sessions in their assigned device, which consisted of flying a visual overhead traffic pattern entry at an uncontrolled airport. Finally, at the end of the second week, all students conducted a post-test flight in the same Frasca ATD as was used during the pre-test. Table 1 shows the experimental procedure used for this study.

Table 1

Experimental Procedure

Group	Assignments	Pre-Test	Training	Post-Test
Control	n = 48	Flight Test in Frasca ATD	No Practice Sessions	Flight Test in Frasca ATD
VR Training	n = 42	Flight Test in Frasca ATD	Three Practice Sessions in VR Simulator	Flight Test in Frasca ATD
PC-based Training	n = 30	Flight Test in Frasca ATD	Three Practice Sessions in PC-based Simulator	Flight Test in Frasca ATD

The Frasca ATD was used due to its ability to replicate the aircraft with a high level of accuracy. One additional benefit is that the Federal Aviation Administration (FAA) accepts this training device as a suitable method for logging training time at all levels of pilot training. Due to the high cost of the physical aircraft, the Frasca ATD was accepted by the research team as a suitable method for evaluating the transfer of pilot skills between the pre-test and post-test flights. Figure 2 shows an image of the Frasca ATD that was used for the pre-test and post-test tasks of this study.

Figure 2. Frasca Aviation Training Device (Frasca International, Inc., 2022)

Equivalence Testing

Prior to conducting the mixed methods study, researchers chose to evaluate the three sample groups to verify equivalency, based on self-reported airplane time, simulator time, and previous VR experience. A one-way ANOVA was performed to compare the effect of four demographic variables (airplane time, simulator time, VR experience, and VR familiarity) on the three assigned quasi-experimental groups (no training, PC, and VR). This step was conducted to establish the equivalency of each group in terms of previous aeronautical experience and VR experience.

Quantitative Study

The purpose of the quantitative study was to determine if Virtual Reality (VR) flight simulators have a significant impact on a pilot's performance versus PC-based flight simulators. The hypothesis was that pilots who train in virtual reality simulators would perform significantly better than pilots who train on PC-based flight simulators and thus would transfer these skills more efficiently to the actual aircraft.

Three variables were analyzed to compare a mean difference between pre-test and post-test performance amongst the participants. These variables were altitude, airspeed, and cross-track distance. The three variables were measured as a difference between the participant's actual performance and their expected (instructed) performance at three points along the procedure. Figure 3 shows the guidance for an overhead traffic pattern entry at an uncontrolled airport. This figure is published in the Airplane Flying Handbook (FAA, 2004) as one method of guidance and expectation for traffic pattern entries. This procedure was used during the study as instructional material for the pre-test and post-test flights.

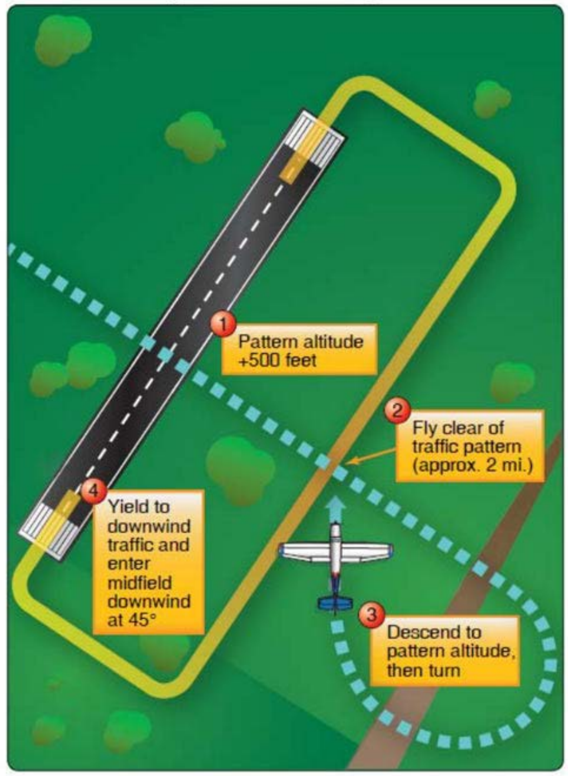

Figure 3. Overhead Traffic Pattern Entry (FAA, 2004).

As stated in the FAA's Airplane Flying Handbook, pilots are expected to cross over the airport at 500 feet above the published traffic pattern elevation (2,400 feet MSL). For this study, an average of the altitude and airspeed parameters were collected as the participants overflew the airport runway in order to gain an accurate measurement during the overfly procedure. The variable of cross-track distance was collected at the physical location where the aircraft crossed the runway and measured as a distance from the midpoint of the runway. The second location where data were collected was abeam the runway touchdown point, where the participant was expected to be at the published traffic pattern elevation (1,900 feet MSL), at an airspeed of 100 knots indicated, and a physical distance of 0.75 nautical miles from the runway centerline. Finally, altitude and airspeed data were captured at a point approximately 0.6 nautical miles from the touchdown point. This point was determined based on a standard 3.0 degree glideslope, where the pilot should be at an altitude of 200 feet above the ground (AGL) and at a stabilized approach speed of 66 knots. Table 2 shows the expected altitudes at each point along the procedure, which were provided to the participants prior to conducting the pre-test and post-test flights.

Table 2

Expected Performance by Variable and Location

	Altitude	*Airspeed*	*Cross-track Distance*
Overfly	2,400’ MSL	100 knots	0.00 NM
Abeam Touchdown	1,900’ MSL	100 knots	0.75 NM
Final at 200 Feet	200’ AGL	66 knots	N/A

Note. MSL = Mean Sea Level. AGL = Above Ground Level. NM = Nautical Miles.

An image of the data collection points is depicted in Figure 4, with a sample flight overlaid on the image. The red circles depict point-in-time locations where data was captured and the two red lines depict the zone where altitude and airspeed data were averaged during the overflight procedure.

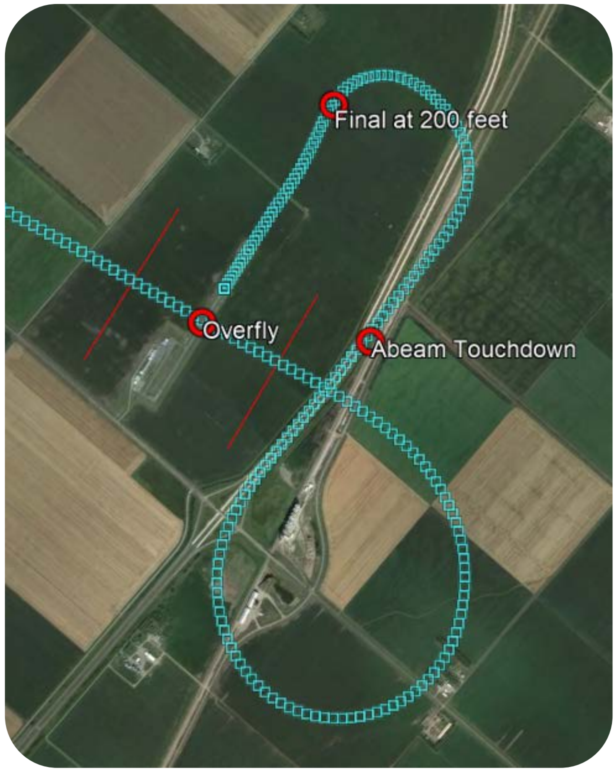

Figure 4. Image of Data Collection Points During the Pre-Test and Post-Test Procedures

Altitude Accuracy Measure

The altitude variable was collected at three points along the procedure. A difference was calculated between the actual flight performance of the participant and the expected altitude for the given location. A final variable of "altitude accuracy" was calculated by taking the absolute value of the difference between the participant's actual flight performance and the expected performance at each of the three points on the procedure. This method allowed researchers to measure the absolute difference of the participant's deviation from the expected parameters across the entire procedure. A small accuracy score would indicate a small deviation from the expected altitude parameters, whereas a high accuracy score would indicate a large deviation from the expected altitude parameters. Altitude accuracy was calculated by using the following equation.

$$
\begin{aligned}
&Altitude\ Accuracy \\
&\quad = \big(ABS(2400 - Overfly\ Altitude)\big) \\
&\quad + \big(ABS(1900 - Abeam\ Touchdown\ Altitude)\big) \\
&\quad + (ABS(200 - Final\ Approach\ Altitude))
\end{aligned}
$$

Airspeed Accuracy Measure

The airspeed variable was collected at three points along the procedure. A difference was calculated between the actual flight performance of the participant and the expected airspeed for the given location. A final variable of “airspeed accuracy” was calculated by taking the absolute value of the difference between the participant’s actual flight performance and the expected performance at each of the three points on the procedure. This method allowed researchers to measure the absolute difference of the participant’s deviation from the expected parameters across the entire procedure. A small accuracy score would indicate a small deviation from the expected airspeed parameters, whereas a high accuracy score would indicate a large deviation from the expected airspeed parameters. Airspeed accuracy was calculated by using the following equation.

$$Airspeed\ Accuracy = (ABS(100 - Overfly\ Airspeed)) + (ABS(100 - Abeam\ Touchdown\ Airspeed)) + (ABS(66 - Final\ Approach\ Airspeed))$$

Cross-track Distance Accuracy Measure

The cross-track distance variable was collected at two points along the procedure. A difference was calculated between the actual flight performance of the participant and the expected position of the aircraft over the ground. A final variable of "cross-track distance accuracy" was calculated by taking the absolute value of the difference between the participant's actual flight performance and the expected performance at each of the two points on the procedure. This method allowed researchers to measure the absolute difference of the participant's deviation from the expected parameters across the entire procedure. A small accuracy score would indicate a small deviation from the expected cross-track distance parameters, whereas a high accuracy score would indicate a large deviation from the expected cross-track distance parameters. Cross-track distance accuracy was calculated by using the following equation.

$$Cross - track\ Distance\ Accuracy = (ABS(0 - Overfly\ Distance\ from\ Midfield)) + (ABS(0.75 - Abeam\ Touchdown\ Distance\ from\ Runway))$$

To evaluate these accuracy parameters, a one-way ANOVA was conducted to compare the effect of each gain score on the assigned quasi-experimental groups. For this step, gain score was calculated by measuring the difference between the post-test and pre-test performance variables to observe the relative performance increase or decrease between the tests. As Huck and McClean (1975) explored, a gain score analysis provides a more accurate picture of the main effect, as we can test for the effect of the treatment (training type) on the pre-test and post-test performance improvement. If we chose a repeated measures ANOVA, as is common for a pre-test/post-test design, the F test for the main effect of treatments would be too conservative, as the treatment only influences the post-test data (Huck & McClean, 1975, p. 512). Altitude accuracy, airspeed accuracy, and cross-track distance accuracy were evaluated to determine the efficacy of VR and PC-based simulator training methods.

Survey Instruments

The purpose of the survey instruments in this study is to understand the perceptions of VR technology and its acceptance and adoption by pilots into the flight training environment. Quantitative and qualitative data were collected in the form of two survey instruments, one during the pre-test procedure and one after the post-test was complete.

The pre-test survey asked for self-reported demographics of the participants' training history prior to the study. The answers to these questions provided the researchers with demographics related to total time in airplanes, total time in simulators, virtual reality experience, and virtual reality familiarity. These demographics were used to evaluate the equivalency of the three sample groups before the mixed methods study was conducted.

The post-test questionnaire was conducted immediately after the post-test flight and asked questions regarding the participants' simulator training experience (PC-based or VR) and evaluated their perceptions on the acceptance and adoption of VR technology for pilot training. Questions were asked to understand how the participants felt they performed on the post-test, as well as providing text-based responses to the advantages and disadvantages of VR technology. These qualitative factors provided indications of the advancement of VR flight simulators compared to more traditional fixed-base training simulators.

There are two techniques this project used to enhance credibility. First, each participant in the study was at an equal benchmark in their flight training experience. This was designed to reduce the performance bias of varying levels of professional experience, which was evident in prior research projects identified in the literature review. Second, the quantitative data collection was modeled to replicate the structure of a previous research project conducted by Reweti, Gilbey, & Jeffrey in 2017. This research project used an identical experimental plan and Frasca ATD to evaluate the participant's level of proficiency during a visually-based airport entry procedure.

Results

Equivalence Testing

A one-way ANOVA was performed to compare the effect of demographic variables on the three assigned quasi-experimental groups. The one-way ANOVA revealed that there was not a statistically significant difference in any of the four demographic variables (airplane time, simulator time, VR experience, or VR familiarity) between any of the three assigned quasi-experimental groups (no training, PC, or VR). Table 3 shows the results of the one-way ANOVA of demographic variables.

Table 3

One-way ANOVA of Demographic Variables

		Sum of Squares	*df*	*Mean Square*	*F*	*Sig.*
Total Time, Airplanes	Between Groups	381.729	2	190.864	.167	.846
	Within Groups	133665.300	117	1142.438		
	Total	134047.029	119			
Total Time, Simulators	Between Groups	100.642	2	50.321	.374	.689
	Within Groups	15759.410	117	134.696		
	Total	15860.052	119			
VR Experience	Between Groups	.274	2	.137	.691	.503
	Within Groups	23.193	117	.198		
	Total	23.467	119			
VR Familiarity	Between Groups	1.111	2	.556	.530	.590
	Within Groups	122.755	117	1.049		
	Total	123.867	119			

Pre-Test versus Post-Test Performance Gain Measures

A one-way ANOVA was performed to compare the effect of each performance gain measure (altitude accuracy, airspeed accuracy, and cross-track distance accuracy) on the three assigned quasi-experimental groups. Additionally, a significant ANOVA result was further analyzed with a post hoc Tukey HSD test for multiple comparisons to determine the effect of mean differences between groups. This post hoc test allowed the researchers to determine which groups showed differences on each of the performance gain measures.

The one-way ANOVA revealed that there was a statistically significant difference in altitude accuracy gain between at least two groups ($F(2,117) = 7.277$, $p=.001$). For airspeed accuracy gain, the one-way ANOVA revealed that there was not a statistically significant difference between at least two groups ($F(2,117) = 0.325$, $p=.723$). Finally, for cross-track distance accuracy gain, the one-way ANOVA revealed that there was a statistically significant difference between at least two groups ($F(2,117) = 21.973$, $p=.001$).

Tukey's HSD Test for multiple comparisons found that the mean value of altitude accuracy gain was significantly different between the PC-based training group and the control group ($p=.033$, 95% C.I. = [10.74, 310.34]), as well as between the VR training group and the control group ($p=.001$, 95% C.I. = [73.26, 345.25]). Additionally, Tukey's HSD found that the mean value of cross-track distance accuracy gain was significantly different between the PC-based training group and the control group ($p=.001$, 95% C.I. =

[0.15, 0.37]), as well as between the VR training group and the control group (p=.001, 95% C.I. = [0.13, 0.33]).

The results of the one-way ANOVA can be found Table 4 and the results of the post hoc Tukey HSD test can be found in Table 5. Additionally, to better visualize the pre-test and post-test results, including the gain score calculations, interaction plots have been provided in Figure 5.

Table 4

One-way ANOVA of Performance Gain Measures

		Sum of Squares	*df*	*Mean Square*	*F*	*Sig.*
Altitude Accuracy Gain	Between Groups	1069840.179	2	534920.089	7.277	.001
	Within Groups	8600956.488	117	73512.449		
	Total	9670796.667	119			
Airspeed Accuracy Gain	Between Groups	63.339	2	31.670	.325	.723
	Within Groups	11400.986	117	97.444		
	Total	11464.325	119			
Cross-Track Distance Accuracy Gain	Between Groups	1.747	2	.874	21.973	.001
	Within Groups	4.651	117	.040		
	Total	6.399	119			

Table 5

Tukey's HSD Test for Multiple Comparisons

	Control		PC-Based		VR		Tukey HSD
	M	*SD*	*M*	*SD*	*M*	*SD*	*p<.05*
Altitude Accuracy Gain	-13.54	214.16	147.00	88.98	195.71	389.36	PC>Control VR>Control PC=VR
Airspeed Accuracy Gain	0.92	8.46	1.60	8.77	2.60	11.90	PC=Control VR=Control PC=VR
Cross-track Distance Accuracy Gain	-0.07	0.14	0.19	0.30	0.17	0.17	PC>Control VR>Control PC=VR

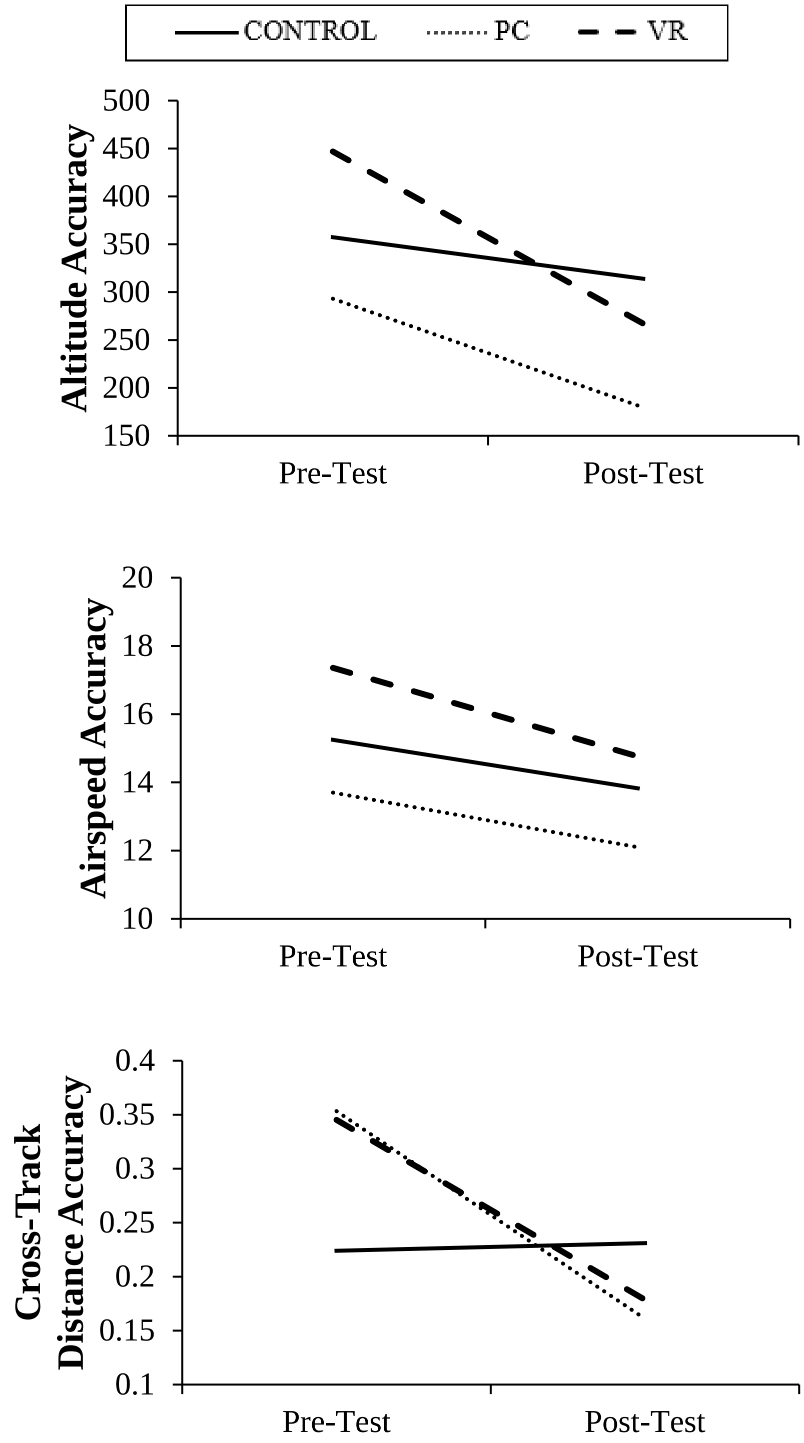

Figure 5. Interaction Plots of Pre-Test versus Post-Test Performance Accuracy

Virtual Reality Acceptance and Adoption

In the post-test survey, participants were asked to rank their performance between the pre-test and post-test flights. The answers to this question were on a three point scale, with options of "worse than the pre-test" (1), "about the same as the pre-test" (2), and "better than the pre-test"(3). A second question asked the participants to compare their training method against existing high-cost flight training devices that are typically used in flight training. This question was asked on a five-point Likert scale, with options ranging from "much worse" (1) to "much better" (5). Finally, participants were asked to provide their thoughts in a series of three open-ended questions. These questions asked the participants to list the advantages, the disadvantages, and any additional thoughts of using VR technology in pilot training. Results of this post-test survey instrument are important to highlight the acceptance and adoption of VR as a viable technology for pilot training.

Pre-Test versus Post-Test Feeling of Performance

For the measure of pre-test versus post-test feeling of performance, students in all three groups felt like they did slightly better in the post-test than they did in the pre-test. Table 6 shows the means and standard deviations for each of the groups.

Table 6

Responses to Pre-Test versus Post-Test Feeling of Performance

	M	*SD*
Control	2.54	0.58
VR Training	2.45	0.62
PC-based Training	2.53	0.56

Training Method Comparison versus High-Cost Flight Training Devices

For the measure of training comparison to a high-cost flight training device, participants who conducted training in both technologies (VR and PC-based) felt like they performed "about the same" as the high-cost training alternative. Table 7 shows the means and standard deviations for each of the groups.

Table 7

Responses to Training Method Comparison to High-Cost Flight Training Device

	M	*SD*
Control	--	--
VR Training	2.47	0.81
PC-based Training	2.88	0.98

Note. Responses were not considered for the "control" group; no training was conducted.

Advantages of VR Technology

In the post-test survey, participants described the advantages of VR technology as a viable method for training student pilots. Many of the responses were consistent in including the "ability to look around the aircraft by using the wings or other objects as reference points helps when flying in the simulator." Several other participants noted that "VR is more readily available and cheaper for students to use" than the high-cost flight training devices. Finally, a few participants included that "VR is available for home use at a reasonable cost," which can "help with preparing for lessons at home, prior to flying the real aircraft."

Disadvantages of VR Technology

When asked what disadvantages VR technology showed in the flight training environment, students were consistent in their answers. Many students stated that "VR does not allow you to interact with the aircraft instruments and equipment." Additionally, "you cannot see or feel the flight controls, which makes it hard to practice tasks like checklist usage." Finally, students who wore glasses noted that the "VR headset didn't fit well and was sometimes blurry when looking through glasses."

Additional Thoughts Regarding VR Technology

The final question in the post-flight survey asked students to provide additional thoughts regarding VR technology and its potential use in the flight training environment. Students overwhelmingly included that "VR simulators are a great learning and practice tool for pilots, as it allows the practice of flows and maneuvers while not being in an actual airplane." Some participants "felt better flying the VR simulators rather than the high-cost flight training devices." A few students said "home simulators and VR significantly helped me progress through my flight training" and "I even noticed that it helped more in my real flight training."

Discussion

The key findings of this study are (1) training in both PC-based and VR simulators provide a significant improvement in the performance of visually-based maneuvers, (2) students believe they perform "about the same" in VR as they do in high-cost flight simulators, (3) students feel that VR is a viable alternative to high-cost flight simulators, and (4) students believe VR simulators help improve their performance in the real airplane. As flight simulation technology improves, flight training organizations will have increased access to low-cost simulator alternatives. This research provides evidence that evaluates pilot performance variables, as well as qualitative acceptance and adoption data, to compare VR training devices and PC-based training devices. These findings will help establish the efficacy of VR technology for pilot training.

The quantitative research in this study addressed the first research question and evaluated three pilot performance variables (altitude, airspeed, and cross-track distance) to compare the pilot's performance before and after their course of training. The results show that PC-based training and VR training significantly increase a pilot's performance on the variables of altitude accuracy and cross-track distance accuracy. For the variable of airspeed, there was no significant effect that training improved a pilot's ability to control that variable. These results can be justified, given that altitude accuracy and cross-track distance accuracy can be more precisely controlled with an immersive visual field. In particular, when learning to fly in the airplane, flight instructors will emphasize the “integrated flight instruction method.” This method teaches pilots to perform flight maneuvers both by outside visual references and by reference to flight instruments (FAA, 2008, p. 9-10). With VR allowing a fully-immersive 360-degree environment, students can more accurately train and control the aircraft with reference to the external environment. Because of this, pilots can better perceive their altitude and cross-track distance, which allows them to make more precise corrections. The variable of airspeed is not as reliant on external visual cues, which is shown in the results of this study, where airspeed was not significantly improved with a training course in either the PC-based or VR simulator.

Qualitatively, this study employed a post-test survey to gauge students’ perceptions on the acceptance and adoption of VR technology for use in pilot training. Research question two asked “how do students’ feel the virtual reality compares to a traditional flight training device?” Overall, students positively responded to the VR technology, stating the “ability to look around the aircraft by using the wings or other objects as reference points helps when flying the simulator.” Additionally, students “felt better flying the VR simulators rather than the high-cost flight training devices.”

Finally, research question three asked, “would students accept virtual reality as a suitable alternative to traditional flight training devices?” The post-test survey revealed that students were somewhat split on this topic. Some students mentioned that “home simulators and VR significantly helped me progress through my flight training” and “noticed that it helped more in my real flight training.” Alternatively, the VR technology did provide some disadvantages, for which students said “it does not allow you to interact with the aircraft instruments and equipment.” Additionally, students who wore glasses mentioned the VR headset was sometimes blurry or foggy, due to the way the headset fit around the glasses. The results of this research question show that there is some room for improvement with VR technology and ergonomics, but students would largely accept virtual reality as a suitable flight training device.

Limitations

One limitation of this study includes the wide variety of available configurations for VR and at-home flight simulators. The study included only one version of a VR flight simulator (running the X-Plane 11 software), which included a VR headset (HTC Vive Pro), force-feedback yoke (Brunner CLS-E MkII), rudder pedals, throttle quadrant, and a trim wheel. The researchers note that low-cost simulator configurations lack the guidance that high-cost fight training devices are required to maintain. Because of this, the results of this study, as well as the perceptions of the VR technology, could change depending on the configuration of the simulator device.

The second limitation noted by the researchers was that of student training outside of this study. Every effort was made to collect research data at the beginning of the academic semester, where students were in the ground training phase of their flight course. This allowed researchers to better control the participants' flight skills, since they were likely not conducting flight training in simulators or aircraft during the period of the research study. That being said, students progress at different rates, as well as have varying levels of access to at-home and personal-use flight simulators. As one participant noted, "I think I got better at the simulator because I flew in the actual airplane a lot during the week of research." This variable was potentially confounding, but outside of the researchers' control, due to the narrow window for data collection.

Finally, as students within the flight training environment are peers and may be placed in different ground training courses, there was a potential limitation of cross-contamination of participant pools. The researchers chose a quasi-experimental method in order to keep all participants conducting each training technology in the same classroom. This allowed for communication and dialogue to center around the exact training method those students were conducting. While this increased the validity of the research, a potential limitation centers around students in one training method conducting an alternate training method, because they learned that a peer in a different classroom was using an alternate method and wanted to try it.

Implication for Practice

The research presented in this paper establishes the grounds for validating the efficacy of virtual reality training devices for pilot training. This research confirms prior literature in flight simulator technology, which shows that training in both a PC-based and VR flight simulator provides significant performance improvements when compared to a control group that received no simulator training. The results of this research, combined with prior literature on the subject, should be considered when attempting to certify and adopt new simulator technologies for use in pilot training.

There is precedent for validating new technology and low-cost aviation training devices in the United States. In the year 2000, the FAA facilitated research to determine which simulator features promote full transfer of pilot performance between simulator and airplane, without unnecessarily driving up cost. The result of this research was the publication of an Advisory Circular titled *FAA Approval of Aviation Training Devices and their Use for Training and Experience* (2018) that defined the requirements for various levels of flight simulation devices, while also establishing the requirement for FAA certification of new simulators that adhere to the defined standard. In particular, the standard established for a basic aviation training device promoted a low-cost option for utilizing approved simulators for initial and recurrent pilot training. This is a substantial improvement that will allow more instructors to utilize the benefits of flight simulators when training student pilots, due to removing the high-cost barrier to entry.

While the FAA has taken steps to understand which simulator features promote skill transfer to the aircraft and have adjusted their policymaking accordingly, there have been significant technological advancements with flight simulators that would benefit from additional research in this domain. In particular, two technologies have seen promising advancements in the flight simulator space, which include the use of virtual reality (VR) and the use of an artificial intelligence-based flight instructor.

Moving forward, additional research could be conducted to evaluate the impact of artificial intelligence-based flight instructor technology on the progress of student pilots. Artificial intelligence-based flight instruction is a technology that provides a student with a pre-determined lesson that covers various skills-based topics. The student will gain an understanding of the skill, receive feedback on their performance while they are flying, and receive an objective score relating to their performance of that skill during the lesson. In this form of simulation, the student would receive similar instruction to what they would receive from an actual flight instructor, which may prove to reduce the risk of primacy, as discussed above. Additionally, this technology could prove useful in allowing students the opportunity to conduct lessons before they fly in the actual aircraft. This benefit would potentially allow student pilots to conduct the lesson in the aircraft more accurately and allow the student to progress through their flight training more efficiently (with less total time to obtain their certificate or rating).

References

Bradley, D., & Abelson, S. (1995). *Desktop flight simulators: Simulation fidelity and pilot performance.* Behavior Research Methods, Instruments, & Computers., 152-159. Psychonomic Society, Inc.

Duncker, K. (1938). *Induced motion.* In W.D. Ellis (Ed.), A source book of Gestalt psychology (pp. 161-172). New York: Humanities Press.

Federal Aviation Administration. (2018). *FAA Approval of Aviation Training Devices and Their Use for Training and Experience* (Advisory Circular 61-136B). Washington, DC: FAA, Department of Transportation.

Frasca International, Inc. (2022). http://www.frasca.com/fixed-wing-simulator-products/

Go, T. H., Burki-Cohen, J., & Soja, N. (2000). *The effect of simulator motion on pilot training and evaluation.* AIAA Modeling and Simulation Technologies Conference.

Huck, S. W., & McLean, R. A. (1975). *Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task.* Psychological Bulletin, 82(4), 511–518. https://doi.org/10.1037/h0076767

Lutte, R.K., & Lovelace, K. (2016). Airline Pilot Supply in the US: Factors Influencing the Collegiate Pilot Pipeline.

McElhiney, A. (n.d.). *Link trainer flight simulator.* Naval Air Station Fort Lauderdale Museum. Retrieved November 28, 2021, from https://www.nasflmuseum.com/link-trainer.html.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2019). *A systematic review of immersive virtual reality applications for higher education: Design Elements, lessons learned, and research agenda.* Computers & Education. Retrieved from https://www.sciencedirect.com/science/article/pii/S0360131519303276

Reweti, S., Gilbey, A., & Jeffrey, L. (2017). *Efficacy of low-cost pc-based aviation training devices.* Journal of Information Technology Education: Research, 16, 127-142.

Rock, I., & Victor, J. (1964). *Vision and touch: An experimentally created conflict between the senses.* Science, 143, 594-596.

United States. Federal Aviation Administration. (2004). *Airplane Flying Handbook*. U.S. Dept. of Transportation, Federal Aviation Administration.

United States. Flight Standards Service. (2008). *Aviation instructor's handbook.* U.S. Dept. of Transportation, Federal Aviation Administration, Flight Standards Service.

Study 3

(As Prepared for Publication)

QUANTIFYING THE TRANSFER EFFECTIVENESS OF AN ARTIFICIAL INTELLIGENCE-BASED SIMULATOR PRE-TRAINING PROGRAM FOR STUDENT PILOTS

Ryan P. Guthridge

University of North Dakota, Grand Forks

Abstract

Since the airline pilot shortage was initially studied in 2016, the pilot hiring model has been significantly impacted, with airlines hiring qualified pilots at unprecedented rates. The COVID-19 pandemic has slowed this hiring rate, however it is expected that airline hiring will soon increase to a rate higher than initially expected (Bureau of Transportation Statistics, 2022). With this dynamic, certified flight instructors are often the most qualified recruits for airlines, due to the number of hours and experience they have gained in the flight training organization. In turn, certified flight instructors are in short supply for flight training organizations worldwide. This study explores a solution to help flight training organizations increase the proficiency of their new student pilots and increase the efficiency of their students' flight training progress as they earn their certificates and ratings. To address these concerns, an artificial intelligence-based technology was evaluated that provided a simulator pre-training program for student pilots (n = 37) prior to beginning their Private Pilot training. The two one-sided test (TOST) procedure was used to evaluate the equivalence of the training and control groups. Then, Roscoe's Transfer Effectiveness Ratio was used to evaluate the effect of a simulator pre-training program on the pre-solo training outcomes of student pilots. The results showed that a guided simulator pre-training program provides a reduction in flight training hours, ground training hours, and the number of calendar days required to complete their pre-solo block of training. These results will help inform flight training organizations who are considering new ways to help support their training pipeline and increase the training efficiency of their organization.

Keywords: artificial intelligence, flight instructor shortage, flight training, pilot shortage, simulator pre-training, student pilots

Literature Review

Aviation, much like any academic discipline, benefits from the use of technology to assist an instructor in delivering content. Similar to a class in the laboratory sciences, Aviation provides a two-part model of instruction. The students must commonly attend a ground school class to learn the knowledge-based topics, while also conducting a laboratory course that teaches them the skill-based maneuvers that are required to earn their certificate or rating. To support this training model, various training technologies are used to help increase student skills at a lower cost than operating an actual aircraft.

One training technology that is widely-used and widely-accepted training in the Aviation industry is the flight simulator. Flight simulators come in many forms, including personal computer based (PC-based) simulators, fixed based simulators (including aviation training devices and flight training devices), and full-motion flight simulators. While each of these simulation options can be obtained at varying levels of cost, they are all considered low-cost when compared to the cost of flying in an actual aircraft. Flight training is costly and simulators provide a low-cost alternative to acquire the skills required for both initial and recurrent flight training.

Broadly, flight simulators have become a widely-accepted training technology that has been proven to develop a student’s skills-based performance, without the requirement to operate a physical aircraft. This is a huge advantage when considering the costs and risks associated with training new pilots in a large aircraft, especially when that aircraft primarily operates to carry paying passengers during commercial air service flights. Most recently, flight simulators have become a foundational technology to provide recurrent training for pilots beginning to fly after long delays due to the COVID-19 pandemic. The risk is profound, especially when evaluating the reports of safety incidents related to a lack of flying due to the widespread ground of aircraft during the COVID-19 pandemic. Figure 1 shows the number of safety incidents reported through the FAA's Aviation Safety Reporting System (ASRS), which describes a consistent risk of incidents related to a lack of flying during the COVID-19 pandemic. On average, nine reports were submitted per month in the year after the pandemic. As a comparison, there were only two reports submitted related to a lack of flying in the years 2018 and 2019 combined.

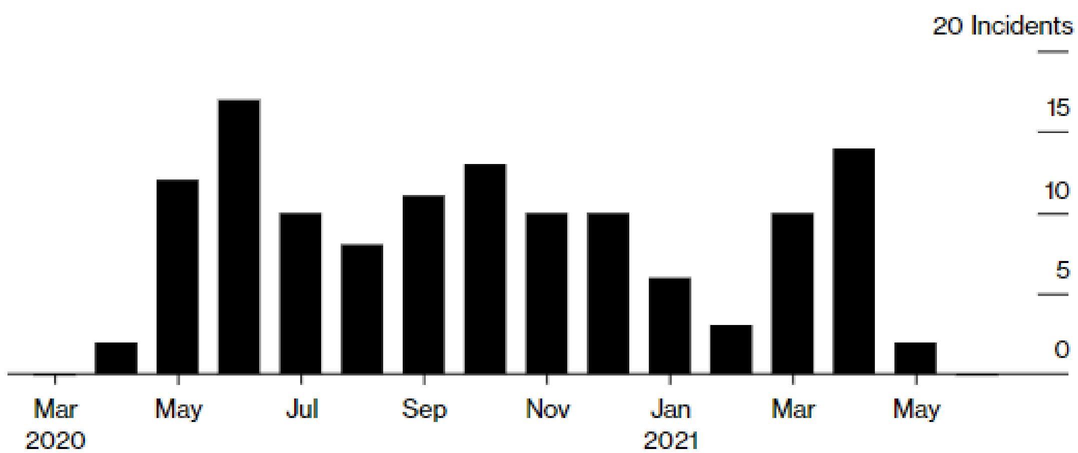

Figure 1. Mistakes by out-of-practice pilots during the COVID-19 pandemic (Whitley & Kotoky, 2021).

As a result of these findings, various airlines have implemented more rigorous recurrent training curricula that use flight simulators as a foundational technology to regain the skills required to operate an aircraft safely. For example, Sidney-based Qantas Airlines puts its Boeing B737 pilots through a six-day course before they get back into the sky, and a senior training captain sits in on their initial flights. The airline's Airbus A380 pilots have two days of training on the ground and in the simulator every 90 days (Whitley & Kotoky, 2021). This is more rigorous than what is required by regulators, which is commonly a routine recurrent training curriculum every 6, 9, or 12 months, depending on the timeline established by the airline and the regulator (Federal Register, n.d.).

The technology in flight simulators has sufficiently advanced to the point where regulators around the world have embraced their use to train pilots during their initial training, as well as during their recurrent training programs. That being said, one key factor that supplements a flight simulator is that of guided instruction. To this point, certified flight instructors have provided that guidance, however as technology develops, there may be viability to utilizing artificial intelligence to provide the required guidance to students.

The Law of Primacy

Perhaps the biggest risk when considering the use of flight simulation technologies is the concept of negative transfer of learning from the simulator to the aircraft. The FAA's Aviation Instructor's Handbook describes the *law of primacy* as one of the most significant factors in the long-term retention of knowledge and skills-based information. They provide the following example:

> When an error occurs pouring a concrete foundation for a building, undoing and correcting the job becomes much more difficult than doing it right the first time. Primacy in teaching and learning, what is learned first, often creates a strong, almost unshakable impression and underlies the reason an instructor needs to teach correctly the first time. (FAA, 2008, p. 3-13)

The risk of violating the law of primacy exists when using flight simulation technologies without guided instruction. In today's environment, flight simulators are easily accessible to students through various forms. These may include the flight school's simulators, a PC-based simulator, or a flight simulator application that they download on their mobile phone or tablet. This level of access enables students to more easily practice outside of the flight training environment, however it introduces the risk of learning skills incorrectly when conducted without the guidance of a flight instructor.

When considering the law of primacy instructors should be careful to ensure that the student learns the skill correctly during the first attempt. When a student learns tasks in isolation, the skill is not initially applied to the overall performance. Additionally, if the skill needs to be relearned, the process can be confusing and time consuming. "The first experience should be positive, functional, and lay the correct foundation for all that is to follow" (United States, 2008). To protect against this negative transfer of learning, instructors should carefully monitor students and guide them to appropriate flight simulation training aids that will improve their performance in the airplane. This could include physically being there to instruct the student during the first attempt at learning a new skill, or the instructor could employ artificial intelligence technologies that have been developed in recent years to help guide the student.

The Artificial Flight Instructor

To help address the concern of negative transfer of learning, a small set of companies have leveraged the widespread availability of flight simulators to develop a technology that uses artificial intelligence to guide students through the primary tasks of flying an airplane. These two companies have advertised the technology as a "training supplement designed to help you achieve your goals faster and for less money" (Redbird Flight Simulations, 2021). To utilize this capability, a student would install a program on their flight training device that contains a pre-configured set of training lessons. As they conduct the training, an artificial intelligence-driven algorithm will instruct the student as they fly the maneuver. It first describes how to conduct a particular skill, demonstrates the skill, and then uses audio to guide the student through the skill as they are flying. This is a similar method to how a flight instructor would teach in the airplane. After the lesson, the student receives a quantitative score on how well they performed the skill, however the student lacks the ability to receive real-time subjective feedback on their performance through a debrief with an instructor. While this is a significant advancement in mitigating the risks of primacy when learning in a flight simulator, more research is needed to confirm the skill transfer effectiveness of this style of flight simulator training.

Roscoe's Transfer Effectiveness Ratio (TER)

In a 2003 study conducted by Taylor and colleagues, flight training devices (FTD) were evaluated to investigate the transfer effectiveness of various amounts of training in the simulators. The research team modeled their study after an early research effort that evaluated PC-based aviation training devices (PCATD) at various increments of training. Both studies used an industry-accepted formula for evaluating the transfer effectiveness of training in flight simulators. Developed by Roscoe (1971), the "transfer effectiveness ratio (TER) accounts for the amount of prior training in ground trainers by specifying the trials/time saved in the airplane as a function of the prior trials/time in the ground trainer." In the Taylor and colleagues (2003) study, four levels of time in a flight training device (5, 10, and 15 hours, respectively) were compared against similar groups in a prior study that used a PCATD. In both studies, the greatest transfer effectiveness ratio was found in the training group that received five hours of prior training in a flight training device, with successively less incremental effectiveness as the training increased to 10 and 15 hours.

This phenomenon was observed and evaluated in a study by Povenmire and Roscoe (1973) which indicated that incremental transfer effectiveness is a negatively decelerating function. This means that with each incremental increase in training using a flight simulator past the optimum threshold, less effectiveness is observed in the actual aircraft. In the Povenmire and Roscoe (1973) study, three levels of training groups were compared with a control group, which received no training in a flight simulator. The simulator used was a generic single-engine flight simulator that did not match the configuration of the aircraft used to evaluate the transfer effectiveness. The study found that each of the groups required less time to pass the flight check than the control group required. “The control group required 45.42 hours, while the transfer groups required 40.26, 38.62, and 37.93 hours, respectively” (Povenmire & Roscoe, 1973).

In both studies, the effectiveness of flight simulators to transfer skill-based training to the actual aircraft was confirmed. All training groups in a generic simulator, PCATD, and flight training device proved that simulators were successful in transferring skills to the actual aircraft. In the Povemmire and Roscoe (1973) study, the TER was 0.3 for the group that spent 7 hours training in the flight simulator. This ratio indicates that for each 10 hours of simulator training, flight training time is reduced by 3 hours, on average.

These studies have influenced FAA regulation, which now requires simulators to be authorized by the FAA prior to their use in flight training, and now permits a significant portion of initial flight training to be completed in an approved PCATD, which may, or may not, replicate the actual aircraft used for flight training. Additionally, the FAA has allowed the use of simulators for instrument pilot currency, without direct oversight from a flight instructor (FAA, 2018). This shows the FAA’s acceptance of flight simulation technologies as a tool that may be used by student pilots to advance their skills, without direct interaction from a human flight instructor.

Purpose of Study

The purpose of this research study is to explore how an artificial intelligence guided simulator pre-training curriculum affects a student pilot's performance in the pre-solo block of the Private Pilot curriculum. The student pilots selected to conduct the training were enrolled in a collegiate aviation degree program at a prominent midwestern University in the United States. All student pilots that participated in this study were unable to immediately enroll in the Private Pilot curriculum their first semester. Half of the students were assigned a guided simulator pre-training curriculum and half conducted no simulator training during their first semester. When the student pilots began the Private Pilot curriculum in their second semester, observations were made to compare student pilots in the simulator pre-training group with the group of student pilots who conducted no simulator pre-training. Analyses were conducted to compare flight training hours, ground training hours, number of lessons to complete the pre-solo block of training, and the number of calendar days to complete the pre-solo block of training.

The technology used for the artificial intelligence guidance was a novel solution provided by TakeFlight Interactive (2022). The artificial intelligence technology was combined with virtual reality simulators, contained in a laboratory available to student pilots at the flight school. This technological setup is a low-cost and high-value option that could be deployed at any flight school to conduct a simulator pre-training program prior to enrolling in a Private Pilot curriculum. The guided simulator lessons are self-paced and can be completed anytime. Issues with negative transfer of learning are reduced with the introduction of artificial intelligence-based instruction, which advises the student pilots of techniques and methods for achieving successful performance of the flight maneuvers.

The results of this research will be used to inform the aviation training industry of the effectiveness of a guided simulator pre-training program for student pilots. The research will inform flight schools of the value of artificial intelligence technology and deploying a simulator training program prior to enrolling in the Private Pilot curriculum. Additionally, the research will be effective in identifying the impact of a guided pre-training curriculum for flight training organizations that may be lacking qualified flight instructors to teach new student pilots.

Methods

The primary outcome of this research project is to evaluate how a guided simulator pre-training program can be used to enhance pilot training. A quantitative approach was used to assess a student pilot's task performance after training in a simulator device using artificial intelligence-based guided lessons. The hypothesis was that student pilots who conduct a guided simulator pre-training program would perform significantly better during the pre-solo block of their Private Pilot training than student pilots who conducted no simulator pre-training. The two groups of participants were evaluated for equivalence using demographic data that was self-reported during a participant qualification survey. This survey asked questions about the student's age, high school GPA, and prior flight training experience. Then, Roscoe's Transfer Equivalence

Ratio (TER) was conducted on each group to evaluate the transfer effectiveness of the simulator pre-training program on the outcomes of the pre-solo block of Private Pilot training.

The Two One-Sided Test (TOST) Procedure for Equivalence

For this study, researchers chose to evaluate the two sample groups to verify equivalency, based on self-reported demographic parameters. The equivalence test was used to statistically reject effects large enough to be deemed worthwhile (Lakens, 2017). The two one-sided test procedure was used to establish an upper and lower equivalence bound, which considered the researcher's smallest effect size of interest. The two composite null hypotheses were tested for each of the one-sided tails. The researchers looked for evidence that the observed effect fell within the equivalence bounds and was close enough to zero to be practically equivalent (Seaman & Serlin, 1998).

In this research project, researchers chose to evaluate an effect size equal to one standard deviation from the mean for each of the demographic variables. The entire sample population of 37 student pilots were used to evaluate these demographics, based on the expected end-state pooled sample population once all participants completed their training.

Quantitative Study

The purpose of the quantitative study was to determine if an artificial intelligence-based simulator pre-training program would positively impact the flight training progress of Private Pilot students in the pre-solo block of training. The participants in this study were selected from students in the Introduction to Aviation course at a midwestern University in the United States. A sample of student pilots were taken from the total, all of who conducted no training their first semester and waited to begin their Private Pilot training their second semester at the University. From this sample, half of the student pilots conducted no simulator pre-training and the other half conducted a guided, artificial intelligence-based, simulator pre-training curriculum during their first semester.

Student pilots who conducted no simulator pre-training were considered the control group and were used as a baseline to evaluate the impact of the simulator pre-training program. Student pilots who were assigned the simulator pre-training course were considered the training group and conducted a self-paced curriculum, which provided guidance from an artificial intelligence-based flight instructor based on their real-time performance in the simulator.

The purpose of these methods, and of using the artificial intelligence-based instructor, is to evaluate an environment by which the student pilot conducts the simulator training entirely independent from the flight training organization. Prior research has been conducted on the transfer effectiveness of a PC-based flight simulator, with the guidance of a human flight instructor (Povenmire & Roscoe, 1973). However, no studies have been conducted on the equivalence of using a similar flight simulator, with the guidance of an artificial intelligence-based flight instructor. The study evaluated a commercial, off-the-shelf solution that a student pilot could reasonably purchase to increase their skills prior to beginning flight training. This solution was developed by

TakeFlight Interactive and was provided at no-cost for the purposes of this study. Additionally, while the study evaluated one type of artificial intelligence-based flight instructor, the company (TakeFlight Interactive) provided no external guidance, review, or financial support to the development of this research paper. All data was collected, processed, and analyzed solely to evaluate the artificial intelligence flight instructor technology and not to evaluate the performance and user experience of the particular software product.

Results

The first step the researchers took during the data analysis phase, was to establish equivalence between the control group and the training group. In total, 37 student pilots participated in the study, with 11 students (29.7%) assigned to the training group (assigned a simulator pre-training course) and 26 students (70.3%) assigned to the control group (assigned no simulator pre-training). The researchers chose to use the two one-sided test (TOST) procedure to evaluate the equivalence of the groups, to ensure the demographics of the two groups could be deemed comparable for the research study. Table 1 shows the mean and standard deviation for the pooled sample population, which was used to determine the upper and lower equivalence bounds for the procedure.

Table 1

Descriptive Statistics for Pooled Population Demographics

	M	*SD*
Age	19.8	3.79
High School GPA	3.46	0.42
Previous Flight Training (hours)	27.90	38.9

Equivalence of the groups was evaluated using the TOST procedure. Additionally, an independent samples *t*-test (null hypothesis significance test) was conducted on each variable to determine the presence of a statistically significant difference between the groups. For all three demographic variables, equivalence of the groups could be verified and no statistically significant difference was found between the groups (Appendix). Taken together, the findings indicate the two groups were statistically equivalent in terms of age, high school GPA, and previous flight training hours (i.e., there were no a priori differences for these variables). Table 2 shows the sample group means, as well as the results of the null hypothesis significance test (NHST) and the two one-sided test (TOST) procedure.

Table 2

Equivalence of Sample Populations

	Control Group *(M)*	Training Group *(M)*	*NHST (p)*	*TOST (p)*
Age	20.0	19.2	.420	.002
High School GPA	3.46	3.44	.909	.014
Previous Flight Training	28.61	26.24	.879	.012

Roscoe's Transfer Effectiveness Ratio (TER)

Since equivalence was verified between the control group and training group, the researchers employed Roscoe's (1971) Transfer Effectiveness Ratio (TER) to evaluate the effectiveness of an artificial intelligence-based simulator pre-training curriculum. In the study, participants were asked to conduct as many simulator training lessons as possible, while following the self-paced guidance of the artificial intelligence-based instructor. This self-paced guidance allowed the student pilots to continue through the training lessons at their own pace, while maintaining a consistent lesson pattern across the sample population.

During the Fall 2021 semester, the training group conducted an average of 8.25 lessons per person, equating to an average of 4.125 hours of simulator training time. Comparably, the control group conducted no simulator training during Fall 2021 the semester. Both the control group and the training group began their initial Private Pilot training in the subsequent semester, Spring 2022. Due to the nature of flight training, a comparably small proportion of the total population was able to complete pre-solo block of training. Because of this, the study was able to examine 4 participants in the training group and 5 participants in the control group by the time of this publication.

Four benchmarks were measured for each student pilot as they conducted the pre-solo block of their Private Pilot training. The pre-solo block of training in this particular Private Pilot curriculum contains the initial pre-solo tasks, which serve as a foundation for developing piloting skills in the aircraft. As such, the researchers chose to examine the transfer effectiveness of the simulator pre-training curriculum against common benchmarks used for training progress and efficiency. These metrics were Flight Training Hours, Ground Training Hours, Number of Lessons required to complete the pre-solo block of training, and the number of Calendar Days between Lesson 1 and Lesson 12 (the final lesson in the pre-solo block). Flight Training Hours and Ground Training Hours were considered metrics for training progress, or total workload expended by the student pilot and flight instructor. Alternatively, Number of Lessons and number of Calendar Days were considered metrics for training efficiency, or how proficient the student pilot was during training and how quickly the student pilot progressed through the curriculum.

Roscoe's TER compares the difference between progress of the control group and the training group, divided by the hours (or units) of training conducted by the training group. The ratio is calculated by using the formula below, where *C* represents participants in the control group and *T* represents participants in the training group.

$$TER = \frac{C_{actual} - T_{actual}}{T_{training}}$$

Typically, the TER is calculated with units of time (hours) of training and time (hours) of actual progress. For instance, in the Povemmire and Roscoe (1973) study mentioned above, the TER was 0.3 for the group that spent a moderate amount of time in the simulator prior to beginning their flight training.

For this study, researchers chose to apply the TER in the same method as Povenmire and Roscoe (1973) to compare the effectiveness of their PC-based simulator training (conducted with an actual flight instructor present) with the artificial intelligence-based flight instructor of this study. Additionally, the researchers chose to apply the TER method to the other three variables of Ground Training Hours, Number of Lessons, and Calendar Days, to evaluate how a structured simulator pre-training program could benefit the student pilot's overall proficiency and efficiency when they enter the flight training environment.

When considering the transfer effectiveness of an artificial intelligence-based simulator pre-training program on the total flight hours realized by initial student pilots in the Private Pilot curriculum, the TER was 0.34 for student pilots who conducted simulator pre-training. This ratio indicates that for each 10 hours of simulator pre-training, flight training time is reduced by 3.4 hours, on average. This compares very closely with the results of Povenmire and Roscoe (1973), who found a TER of 0.3 for student pilots who conducted training in the simulator.

For the other metrics gathered in this study, Ground Training Hours resulted in a TER of 0.09. This indicates that for each 10 hours of simulator pre-training, ground training time is reduced by 0.9 hours. The number of lessons it took for student pilots to complete pre-solo block of training resulted in a TER of -0.02, which means that for each 10 hours of simulator pre-training, the number of lessons is increased by 0.2. Finally, the number of calendar days it took to complete the pre-solo block of training resulted in a TER of 5.12. This indicates that for each 10 hours of simulator pre-training, the number of calendar days it takes to complete the pre-solo block of training is reduced by 51.2 days. Table 3 contains a complete list of sample means and associated TER values.

Table 3

Transfer Effectiveness Ratio (TER) of Factors in Pre-Solo Training

	Control Group *(M)*	Training Group *(M)*	*TER*
Flight Training Hours	18.8	17.4	0.34
Ground Training Hours	4.2	3.9	0.09
Number of Lessons	18.4	18.5	-0.02
Calendar Days	80.6	59.5	5.12

Note: The TER value of $T_{training}$ equals 4.125 hours for the variables in this table.

Discussion

The key findings of this study are that an artificial intelligence-based simulator pre-training program (1) significantly reduces the number of calendar days to complete the pre-solo block of Private Pilot training, (2) reduces the flight training hours required in the pre-solo block of Private Pilot training, (3) reduces the ground training hours required in the pre-solo block of Private Pilot training, and (4) has little impact on the number of lessons required to complete the pre-solo block of Private Pilot training. As the Aviation industry climbs out of the COVID-19 pandemic and hires airline employees at pre-pandemic rates (Bureau of Transportation Statistics, 2022), flight training organizations will be challenged with finding solutions to combat a shortage of certified flight instructors at their flight school. The trend in the aviation industry will quickly return to a point where the airlines are hiring qualified flight instructors faster than the civilian and military sectors can produce newly-qualified pilots to take their place. These dynamics influence the rate at which student pilots complete their training and alternative solutions should be considered to support the production of new pilots in the flight training environment. The findings in this study are particularly important to highlight the advancing technology of the artificial flight instructor and quantify the impact of a self-paced simulator pre-training program for student pilots.

Flight Training Hours

When enrolled in the initial Private Pilot curriculum, student pilots spend the pre-solo block of training learning foundational skills required to maneuver the aircraft and navigate in the national airspace system. On average, student pilots spend approximately 19.0 hours in the aircraft learning these skills. This study evaluated the effect of employing an artificial intelligence-based simulator pre-training curriculum to student pilots who were unable to immediately enroll in the Private Pilot curriculum their first semester at the University. This allowed the student pilots one full semester to conduct simulator pre-training lessons, with the support of artificial intelligence to guide the student pilot through the maneuvers and lessons.

The study found that the transfer effectiveness ratio (TER) of an artificial intelligence-based simulator pre-training program on the total flight hours realized by initial student pilots in the Private Pilot curriculum was 0.34 for student pilots who conducted simulator pre-training. This ratio indicates that for each 10 hours of simulator pre-training, flight training time is reduced by 3.4 hours, on average. This compares very closely with the results of Povenmire and Roscoe (1973), who found a TER of 0.3 for student pilots who conducted training in the simulator.

The findings of the study are notable, considering the similarities between the TER for an artificial intelligence-based instructor and the TER of the Povenmire and Roscoe (1973) study, which utilized the guidance of a human flight instructor. While this is one metric to consider when attempting to employ a new technology for student pilot training, it is important to understand the relevance of the results. With an artificial intelligence-based flight instructor performing comparably to a human flight instructor in respect to the transfer effectiveness of simulator pre-training, all student pilots could benefit from conducting a guided simulator pre-training program prior to enrolling in the Private Pilot course of training. Additionally, flight training organizations who find themselves with a reduced amount of flight instructors to teach student pilots should consider employing an artificial intelligence-based solution to reduce the risk of negative learning transfer, as indicated in the literature review. In either case, student pilots can benefit from instructor-guided simulator pre-training, which has an effect of approximately 3 hours of reduced flight training hours for each 10 hours spend in the simulator.

Ground Training Hours

In addition to evaluating flight training hours, this study also evaluated ground training hours as a measure of student pilot proficiency that resulted from a simulator pre-training course. The hypothesis was that student pilots who conducted a simulator pre-training course would be more proficient on the pre-solo training maneuvers and take less time learning that content with their flight instructor during ground training. An additional hypothesis was that student pilots who conducted a simulator pre-training program would show a higher level of preparedness during lessons, due to the artificial intelligence guidance of the pre-training program. The study showed that ground training hours resulted in a TER of 0.09. This indicates that for each 10 hours of simulator pre-training, ground training time is reduced by 0.9 hours.

The findings of this study reveal that student pilots who conduct a simulator pre-training course spend a moderately-less amount of time in ground training hours versus their non-training counterparts. While 0.9 hours of ground training time may not seem significant, when compared to the average of the total population, the results can be put into context. On average, student pilots in the sample population spent 4.5 hours of ground training time in the pre-solo block of training. To contextualize the findings of the study, student pilots who conducted a simulator pre-training course spent 20 percent less ground training hours than student pilots who conducted no simulator training prior to their Private Pilot course. While ground training content has strict requirements dictated by the FAA, there are no ground training hour requirements to complete this training. As such, these results should continue to be monitored, as they may change depending on the ground training hour requirements of the respective flight school.

Number of Lessons

In this study, the metric with the least amount of effect was the number of lessons required to complete the pre-solo block of training. The number of lessons it took for student pilots to complete the pre-solo block of training resulted in a TER of -0.02, which means that for each 10 hours of simulator pre-training, the number of lessons is increased by 0.2. While one could infer that simulator pre-training results in an increase in the amount of lessons required to complete the pre-solo block of training, the magnitude of this effect is fairly small. On average, the sample population required 19.2 lessons to complete the pre-solo block of training. With an effect of 0.2 lessons, this equates to a one percent increase in lessons required.

Calendar Days

The largest effect shown in the study was the number of calendar days required to complete the pre-solo block of training. In the sample populations, all student pilots began their flight training in January of the Spring 2022 semester. While there are numerous variables that may slow down a student pilot's training progress (weather, holidays, flight instructor and airplane availability), there was consistency in that each student pilot in the study was conducting their training during the same calendar period.

The findings of this study reveal that student pilots who complete a simulator pre-training course spend significantly less time in the pre-solo block of training than those student pilots who conduct no training. In the study, the number of calendar days it took to complete the pre-solo block of training resulted in a TER of 5.12. This indicates that for each 10 hours of simulator pre-training, the number of calendar days it takes to complete the pre-solo block of training is reduced by 51.2 days. This finding is significant in understanding that student pilots who conduct simulator pre-training progress faster through the curriculum than those student pilots who conduct no training. For a flight training organization that is looking to increase the efficiency of their student pilots and are looking for them to progress faster through the training curriculum, this study encourages the use of a simulator pre-training program to achieve that goal.

Limitations

In this study, there are a few limitations the researchers would address. First, while the sample population of 37 participants yielded an appropriate power for the study, there were fewer student pilots that had completed the pre-solo block of training at the time this study was published. Ultimately, nine participants out of the sample population of 37 were able to be analyzed for the transfer effectiveness of the simulator pre-training program. Ideally, while the researchers would prefer to analyze the full sample population, there was an uncontrolled variable at play, in that the student pilots were unable to complete the training course by the time this dissertation publication was approved.

Second, this study employed only one version of an artificial intelligence-based instructional technology. At the time of this publication, there are two companies that produce an artificial intelligence-based instructional technology for aviation pilot training. These companies are TakeFlight Interactive (the technology used for this study) and Redbird Flight Simulations. While it is important to note that both of these companies use a core technology that was created by the same development team, there may be additional artificial intelligence offerings for flight schools to adopt, which may yield different results.

Finally, as noted in the Discussions section of this paper, the metric of "calendar days" is further confounded due to factors outside of a student pilot's control. These external factors include poor weather, holidays, flight instructor availability, airplane availability, and illness, to name a few. While every effort was made to control the validity of this variable with all student pilots beginning their training in the same calendar month (January) of the same year (2022), it is expected that there were unaccounted-for variables to this metric that were outside of the researcher's control.

Implication for Practice

The results of this study show that consideration should be placed in implementing a guided simulator pre-training program for student pilots. This pre-training program should employ a flight instructor guided element, to reduce the risk of a negative transfer of learning. That being said, this study shows a comparable level of transfer effectiveness in utilizing a human flight instructor versus using an artificial intelligence-based flight instructor. As the Aviation industry climbs out of the COVID-19 pandemic and hires airline employees at pre-pandemic rates (Bureau of Transportation Statistics, 2022), flight training organizations should explore utilizing new and advancing artificial intelligence-based technologies to reduce the impact of a shortage of certified flight instructors at their flight school.

Study and research of this topic in the aviation industry is integral to improving the flight training progress of student pilots, regardless of flight instructor shortage concerns within the flight training organization. Beyond the research presented in this paper, it is suggested to expand upon the sample populations and employ these statistical methods at alternative flight training organizations around the world. While it is hypothesized that alternative flight training organizations would yield similar results to this study, it is unknown if varying structures of training organizations would realize the same benefit of a simulator pre-training program, as was found in this study. Regardless, all flight training organizations should consider the use of a guided simulator pre-training program to increase the proficiency, efficiency, and capabilities of student pilots within the training program.

References

Bureau of Transportation Statistics. (2022). U.S. Cargo and Passenger Airlines Add 20,049 Jobs in December 2021, Reaching COVID-19 Pandemic High; Employment Remains 2.9% Below Pre-Pandemic December 2019. United States Department of Transportation. https://www.bts.gov/newsroom/us-cargo-and-passenger-airlines-add-20049-jobs-december-2021-reaching-covid-19-pandemic

Federal Aviation Administration. (2018). *FAA Approval of Aviation Training Devices and Their Use for Training and Experience* (Advisory Circular 61-136B). Washington, DC: FAA, Department of Transportation.

Federal Register. (n.d.). Retrieved July 6, 2022, from https://www.ecfr.gov/current/title-14/chapter-I/subchapter-G/part-121/subpart-N

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, 8(4), 355–362. https://doi.org/10.1177/1948550617697177

Povenmire, H. K. & Roscoe, S. N. (1973). Incremental transfer effectiveness of a ground-based general aviation trainer. *Human Factors, 15, (6),* 534-542.

Redbird Flight Simulations. (2021). Redbird Gift. Redbird GIFT. Retrieved December 2, 2021, from https://gift.redbirdflight.com/.

Roscoe, S. N. (1971). Incremental transfer effectiveness. *Human Factors, 13, (6),* 561-567.

Seaman, M. A., Serlin, R. C. (1998). Equivalence confidence intervals for two-group comparisons of means. Psychological Methods, 3, 403–411. doi:10.1037/1082-989X.3.4.403

TakeFlight Interactive. (2022). https://www.takeflightinteractive.com/

Taylor, H., Talleur, D., Rantanen, E., & Emanuel Jr., T. (2003). *Transfer of training effectiveness of a flight training device (FTD).* FAA Cooperative Agreement 2002-G-033 Semi-Annual Interim Report.

United States. Flight Standards Service. (2008). *Aviation instructor's handbook.* U.S. Dept. of Transportation, Federal Aviation Administration, Flight Standards Service.

Whitley, A., & Kotoky, A. (2021). *Out-of-Practice Airline Pilots Are Making Errors Back in the Air.* Bloomberg.com. Retrieved November 15, 2021, from https://www.bloomberg.com/news/features/2021-10-14/plane-crash-risks-rise-as-covid-19-leaves-airline-pilots-out-of-practice.

Appendix

Equivalence Test Results for Demographic Variables

For the demographic of age, there was no significant effect between the control group and the training group, $t(35) = 0.816$, $p = .420$. The equivalence test was significant, $t(35) = -3.05$, $p = 0.002$, given equivalence bounds of -3.79 and 3.79 (on a raw scale) and an alpha of 0.05. For the demographic of high school GPA, there was no significant effect between the control group and the training group, $t(35) = 0.115$, $p = .909$. The equivalence test was significant, $t(35) = -2.29$, $p = 0.014$, given equivalence bounds of -0.42 and 0.42 (on a raw scale) and an alpha of 0.05. For the demographic of previous flight training (hours), there was no significant effect between the control group and the training group, $t(35) = 0.154$, $p = .879$. The equivalence test was significant, $t(35) = -2.37$, $p = 0.012$, given equivalence bounds of -38.9 and 38.9 (on a raw scale) and an alpha of 0.05.

Conclusion

After reviewing the purpose of this program of research, “to gain an understanding of how flight training organizations can address delayed flight training with the use of advanced training technologies”, some steps can be made to help alleviate the strain of delayed student progress on flight training organizations.

The first step is recognizing the impact of delayed student progress in the flight training environment. As shown in the first study, students who are conducting a flight laboratory that is nonconcurrent to their academic course of study suffer significantly academically. In seven out of eight block exams, students who were in a nonconcurrent flight laboratory scored significantly worse than students in a concurrent laboratory. On average, students in a nonconcurrent laboratory score 5.5% worse on each block exam than students in a concurrent flight laboratory. Functionally, this would be the equivalent to a student receiving a grade of A in the class, versus a student receiving a B+. The disconnect between flight laboratory and ground course progress is shown to be detrimental to a student’s academic success and every effort should be made to avoid nonconcurrent laboratories during their flight training.

The second step to alleviating the strain of delayed student progress is to increase access to affordable flight training devices. The second study in this dissertation evaluated the efficacy of virtual reality, which is a new and largely untested technology in the aviation training space. The study showed that students who conducted training in a virtual reality simulator performed significantly better than a control group who received no simulator training. Comparably, students who trained in a virtual reality simulator showed similar performance improvements to students who trained in a PC-based simulator. Finally, students answered a series of open-ended questions to help researchers understand their perceptions on the acceptance and adoption of virtual reality technology in pilot training. Overwhelmingly, students described the value of virtual reality in preparing for flight lessons. Additionally, they explained the low-cost and high-value benefits that virtual reality provide, which emphasizes the impact that accessible, low-cost flight simulators can provide to student pilots. These findings were significant, as the FAA is exploring what features in simulators promote full transfer of pilot performance between simulator and airplane, without unnecessarily driving up cost. Virtual reality is an advanced technology that can provide a number of benefits in the flight training environment.

The third step to alleviating the strain of delayed student progress is to explore alternative methods for guided flight instruction. As the aviation industry begins to hire flight instructors into commercial pilot roles, flight training organizations will be left with an increasing number of student pilots and few flight instructors available to teach those students. The third study in this dissertation explored the use of an artificial intelligence-based flight instructor to guide student pilots in a simulator pre-training program. Student pilots in this study were assigned a self-paced simulator pre-training curriculum, that was guided with feedback from an artificial intelligence-based flight instructor. The results of the study showed that students who conducted the pre-training lessons (1) took fewer calendar days to complete the pre-solo block of Private Pilot training, (2) required fewer hours in the airplane during the pre-solo block of Private Pilot training, and (3) required fewer ground training hours during the pre-solo block of Private Pilot training. The findings of this study show a comparable level of transfer effectiveness in utilizing a human flight instructor versus using an artificial intelligence-based flight instructor. This is significant when considering methods to supplement reduced flight instructor staffing, while still providing a quantifiable benefit to the student pilots in the flight training organization.

The next steps in this program of research is to expand upon the sample populations of study three and explore the use of artificial intelligence guidance in other areas of the aviation training industry. While it is hypothesized that alternative flight training organizations would yield similar results to this study, it is unknown if varying structures of training organizations would realize the same benefit of a simulator pre-training program, as was found in this dissertation. Additionally, while the efficacy of virtual reality technology was validated in this program of study, alternative uses for virtual reality and augmented reality could prove beneficial in the flight training organization. For instance, augmented reality could be studied as a supplement for aviation maintenance technicians conducting aircraft inspections or as a method to gain a three-dimensional perspective on topics presented in a textbook. Nevertheless, this program of study encourages the use of virtual reality and artificial intelligence technologies to help alleviate the strains of delayed flight progress in aviation pilot training and research should be continued to explore how these technologies can benefit other areas of the aviation industry in the future.