EMBRY-RIDDLE Aeronautical University
SCHOLARLY COMMONS

PhD Dissertations and Master's Theses

Spring 5-2022

Comparing Training Effects of Virtual Reality Flight Simulation to Conventional PC-Based Flight Simulation

Tianxin Zhang
zhangt2@my.erau.edu

Follow this and additional works at: https://commons.erau.edu/edt

Part of the Human Factors Psychology Commons

Scholarly Commons Citation
Zhang, Tianxin, "Comparing Training Effects of Virtual Reality Flight Simulation to Conventional PC-Based Flight Simulation" (2022). *PhD Dissertations and Master's Theses*. 668.
https://commons.erau.edu/edt/668

This Dissertation - Open Access is brought to you for free and open access by Scholarly Commons. It has been accepted for inclusion in PhD Dissertations and Master's Theses by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.

Comparing Training Effects of Virtual Reality Flight Simulation

to Conventional PC-Based Flight Simulation

By

Tianxin Zhang

Master of Science, Embry-Riddle Aeronautical University, 2017

Bachelor of Science, Embry-Riddle Aeronautical University, 2015

A dissertation submitted to the Department of Human Factors and Neurobiology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Human Factors

Embry-Riddle Aeronautical University

Daytona Beach, Florida

Spring 2022

Signature Page

Comparing Training Effect of Virtual Reality Flight Simulation to Conventional PC-Based Flight Simulation

By

Tianxin Zhang

This dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, Dr. Christina M. Frederick, and has been approved by the members of the dissertation committee. It was submitted to the College of Arts and Sciences and was accepted in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Human Factors.

____

Dr. Christina M. Frederick, Ph.D.
Committee Chair

______________________________	______________________________
Dr. Elizabeth L. Blickensderfer, Ph.D. Committee Member	Dr. Scott A. Shappell, Ph. D. Department Chair, Human Factors
______________________________	______________________________
Dr. Barbara S. Chaparro, Ph.D Committee Member	Dr. Karen F. Gaines, Ph.D. Dean, College of Arts and Sciences
______________________________	______________________________
Dr. Andrew R. Dattel, Ph.D. Committee Member	Lon Moeller, J.D. Senior Vice President for Academic Affairs and Provost
	______________________________
	Date

Acknowledgments

I'd first like to sincerely thank my advisor as well as my dissertation committee chair, Dr. Christina M. Frederick, for her mentorship and friendship. She has given me continuous guidance and encouragement with immeasurable patience throughout my time at Human Factors Department. It has been an honor to work with her, and I'll never be able to adequately thank her enough for all that she has done.

I sincerely as well thank my other committee members, Dr. Andrew R. Dattel, Dr. Blickensderfer, and Dr. Barbara S. Chaparro, for their expertise and willingness to be a part of this project. I am very fortunate to have such a diverse and personable group of people to lean on.

Thanks to Mr. Rafael Dubena and Dr. Kenneth P. Byrnes for your comments and suggestions which largely inspired me in the development of the research design.

I'd also like to thank and acknowledge my parents, who are the biggest supporters of my life. My strength and power in pursuing my dreams were attributed to their limitless trust and faith in me.

Abstract

The purpose of the current project was to investigate the effect of utilizing Virtual Reality (VR) technologies for flight training by comparing the training results when using conventional desktop flight simulation versus VR flight simulation. Additionally, this project examined the user experience of VR flight simulation and how users' motivation and satisfaction with VR simulations. This research employed a quasi-transfer of training experiment including 48 participants. Analyses indicated that VR group participants performed better in the post-training maneuver performance on an FTD than in the conventional desktop simulation group. Findings also supported that VR flight simulation could provide a better user experience and generate a higher motivation for usage. This work contributed positive evidence that VR flight simulation has a large potential to be an effective flight training and provided a foundation for future research to continue exploring the training effect of VR flight simulation.

Keywords: virtual reality, flight training, user experience, self-efficacy

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AATD	Advanced Aviation Training Device
ATD	Aviation Training Device
BATD	Basic Aviation Training Device
CGI	Computer Generated Image
DOF	Degrees-of-Freedom
FAA	Federal Aviation Administration
FFS	Full Flight Simulator
FOV	Field of View
FTD	Flight Training Devices
HMD	Head-Mounted Display
IFR	Instrument Flight Rules
ILS	Instrument Landing System
IPC	Instrument Proficiency Check
IVE	Immersive Virtual Environment
PC	Personal Computer
PCATD	Personal Computer Aviation Training Devices
PPL	Private Pilot License
SD	Standard Deviation
TER	Transfer Effectiveness Ratio
UEQ	User Experience Questionnaire
VFR	Visual Flight Rules
VR	Virtual Reality

Introduction

Problem Statement

Development and refinement of pilot skills are critical and time-consuming components of the training curriculum for student pilots. At the same time, flight training imposes a substantial financial cost. The potential to reduce costs using inexpensive but effective training methods is of interest to the aviation community. For the past few decades, the personal computer aviation training devices (PCATD's) served as a low-cost training alternative compared to certified generic flight training devices (FTD's). Studies showed that the training effectiveness of the PCATD's was generally positive and substantial when new tasks were introduced (Homan & Williams, 1997; Taylor et al., 1999; Beckman, 2000; Beckman, 2003).

With the advent of Virtual Reality (VR) technologies, one of the most popular fields of VR application is training. With advantages like high-fidelity, repeatability, flexibility, and low cost (Norris et al., 2019), VR simulation has already been used in medical, military, and mining training, and studies showed that training on VR simulators was valid and transferrable to real environments (Bowman & McMahan, 2007; Chen et al., 2008; Vankipuram et al., 2010; Grabowski & Jankowski 2014; Mayti et al., 2015). While many advantages of VR training exist, research concerning users' experience and acceptance of VR based instruction is still under-explored and not fully known (Chang et al., 2019), and one area in question is whether VR simulation could be used effectively for pilot training,

One concern regarding VR based training is cybersickness. Users of some VR applications have reported experiencing the side effect, known as "cyber-sickness"; it often happens when there is a conflict between the vestibular system and visual perception (Norris, Spicer, & Byrd, 2019). Other side effects of VR can include vertigo, ataxia, disorientation, headache, eyestrain, and nausea (LaViola, 2000). In addition, using VR simulation effectively for training purposes requires the trainee to develop skills to use the technology, know the interface, and develop familiarity with the controls. Thus, in order to assess a training method from a Human Factors point-of-view, beyond studying the learning effectiveness of VR simulation training, it is also important to look at the user experience of VR technologies and users' willingness to adopt such technologies for training.

VR flight simulation seems to have the potential for becoming the new low-cost alternative training method for novice pilots. However, currently, a limited number of studies have focused on evaluating either the learning effect of flight training utilizing VR technologies or the user experience of VR flight simulation. How effective flight training with VR simulation could be and what student pilots' attitudes are toward using VR for training are not clear.

Purpose of Study

The purpose of the current project was to investigate the effect of utilizing VR technologies for flight training by comparing the training results when using conventional PCATD versus VR flight simulation. Additionally, this project examined the user experience of VR flight simulation and how users' motivation and satisfaction are impacted by VR simulations. While the focus of the project using VR simulation for flight training, the ideas, and findings generated by the current project would likely generalize to other areas within VR simulation based training. The aim is that this project will contribute new knowledge and ideas both to the aviation community as well as the broader human factors community.

Literature Review

This chapter discusses a review of related literature concerning the background of flight simulation, simulation-based learning, how the effectiveness of training with flight simulations was measured in existing related studies, the background of VR technologies, and research related to VR simulation for training.

Flight Simulation

History of Flight Simulation

Jones et al. (1985) defined simulation as an interactive system that represents an operational system through artificial duplication or replication of the system and its equipment, environment, and capabilities. Flight simulation refers to technology that reproduces the human-aircraft interaction for the training purpose of performance evaluation, research, and development (Gheorghiu, 2013). The origin of flight simulation can be traced back to the beginning of manned flight. The student pilots of the first powered aircraft were trained by using a low powered machine enabling rudder control to practice taxiing and using a high powered machine with elevator and aileron control to practice (Baarspul, 1990). With these tools, student pilots could learn the feel of basic flight controls, while proceeding along the ground.

The first proposals for full ground-based simulators were based on aircraft attached to the ground, but capable of responding to aerodynamic forces (Baarspul, 1990). The 'Sanders Teacher' was an example of such a device. It was an aircraft mounted on a universal joint in an exposed position and facing into the prevailing wind; it was able to respond in attitude to the elevator, aileron, and rudder controls.

In the 1930s, as technology advanced, relatively more sophisticated electro-mechanical flight simulators became popular (Koonce & Bramble, 1998). The Link Trainer, developed by

American aviator and inventor, Edwin Link, was the most successful and well-known device of this type (Baarspul, 1990). Instrument flying training was started at Links' flying school in the early 1930s; the Link Trainer had full flight controls and instruments, including an artificial horizon, which was crucial to fly instrument. Pitch, yaw, and roll movements were initiated using pneumatic bellows for actuation. The various instruments were operated either mechanically or pneumatically. Flight instructors could watch an external repeater to see how a student manipulated the flight controls in the Link Trainer. In 1937, American Airlines became the first airline to use a Link Trainer for their pilot training. The Link Trainer was produced in many versions and sold in many countries such as England, France, Japan, Germany, and the USSR.

From the 1940s to the 1950s, a major improvement in flight simulation was the use of a type of analog computer to solve the equations of motion of an aircraft (Baarspul, 1990). These analog computers enabled aircraft simulator response to aerodynamic forces as opposed to an empirical reproduction of these effects. However, one main restriction in flight simulation at that time was that aircraft manufacturers did not have much actual flight data on the performance and dynamic characteristics of their airframes and engines. The simulator manufacturers were therefore required to use trial and error methods depending on pilot evaluations, to adjust the simulator so that it would "fly" as much as possible like the aircraft. This situation changed when more and more flight data were gathered to generate a flight simulator database. Many simulators were designed and built with individualized cockpits, controls, and instrument displays for type-specific aircraft training during the 1940s (Baarspul, 1990).

In the early 1960s, Link developed a special purpose digital computer, the Link Mark I (Baarspul, 1990). This computer was first designed for real-time simulation and had three parallel processors for arithmetic, function generation, and radio station selection. By the 1970s, faster and more powerful general purpose computers became suitable for real-time simulation.

Up to the mid-1950s, nearly all the simulators could not provide cockpit motion, which did not allow fixed-base simulators to “fly” like airplanes. In 1969, the first three degrees-of-freedom (DOF; roll, pitch, and heave) motion system using hydraulic actuators with hydrostatic bearings were first introduced by the Faculty of Mechanical Engineering of Delft University of Technology (Baarspul, 1990). The first commercially available 6 DOF (Pitch, Roll, Yaw, Surge, Sway, and argued) motion system was developed in 1977. Later, specialized software, mixed with highly sensitive hydraulic systems, became the core of full-flight simulator.

In terms of visual cues, systems for producing the visual scene for flight have been proposed and constructed since the beginning of flight simulators (Baarspul, 1990). Point-light source projection, film projection, and Closed Circuit Television providing the view of a scale model of an airport by a moving television camera were some early attempts to create a realistic external display. The first computer generated image (CGI) systems for simulation were produced by the General Electric Company for the US space program (Baarspul, 1990). Early versions of these systems were able to generate a two-dimensional (2D) 'pattern ground plane' image, while later versions produced 3D images in real-time. With today’s advances in technology development, CGI systems have increased in performance, speed, and fidelity, providing high quality pictures with accurate real-time visual feedback.

As technologies advanced, and flight simulators achieved higher levels of physical fidelity, the aviation industry gained authorization from the Federal Aviation Administration (FAA) for greater amounts of training in the simulator and less required training in the actual airplane, resulting in greater savings and improved safety in training programs (Koonce & Bramble, 1998).

The 1980s brought advanced in personal desktop computers (PC), and this had implications for flight simulation. In 1981, IBM came out with their PC, bringing the power of flight simulator computers into the reach of the individual general aviation pilot (Koonce & Bramble, 1998). Early PC flight simulation programs were marketed as games, and the flight control inputs were keys of the computer keyboard. Within a few years, control yokes, rudder pedals, and throttles for PC flight simulation became available. Today, with flight simulation software, such as Lockheed Martin's Prepar3D, X-Plane 11, AeroFly2, and Microsoft Flight Simulator X, PC-based flight simulation became a viable tool for presenting graphic representations of aircraft instrumentation, aerodynamic characteristics, and navigations mimicking those experienced in real flight (Talleur et al. 2003). Compare to certified FTDs, PCATDs relatively low-cost and easy-to-access software provided opportunities for student pilots and aspiring pilots to experience simulated flights for learning purposes. Since 1997, the FAA recognized PCATDs as an effective method of obtaining instrument flight training, by allowing ten hours of PCATD training to substitute for aircraft flight time in training for an instrument rating when supervision is given by an authorized instructor (Mcdermott, 2005).

The idea to utilize VR for simulation dated back to the 1960s, when Ivan Sutherland developed the "Ultimate Display" concept that the visual interface should not be thought of as a screen, but rather a window to a virtual world that looks real, sounds real and reacts in real-time (Yavrucu et al., 2011). However, due to the constraints of hardware and software capacity, VR flight simulation was not a feasible or affordable option until fairly recently, after the releases of the Oculus Rift and the HTC VIVE (Oberhauser et al., 2018). With the current hardware and software advances in the field of VR, the VR flight simulation systems with flight simulation software such as Xplane-11, AeroFly2, could evolve traditional ATD into a more sophisticated tool for training (Oberhauser et al., 2018).

Types of Flight Simulations

There are three types of ground training devices that are recognized by the FAA for flight training purposes (Beckman, 2000). The first type of device is called the Full Flight Simulator (FFS). They are multi-million-dollar machines with highly sensitive hydraulics and have full visual displays. There are 4 levels in FFS, Level A to D. Level A has 3 axis motion with night visuals, Level B has 3 axis motion, night visuals, and ground handling simulation, Level C has 6 axis motion, night & dusk visuals, dynamic control loading with higher fidelity, and Level D has 6 axis motion, night, dusk & day visuals with dynamic control loading providing the highest fidelity. The use of FFS is typically limited to airline use or professional pilot training programs, due to both their initial and operating costs (Beckman, 2000). The next type of training device is called the Flight Training Device (FTD). FTD has generally replicated an aircraft cockpit and often has a visual display system but provides no motion feedback. There are seven levels of FTD's, however, levels 1, 2 & 3 are no longer issued (FAA, 2018). Level 4 is a basic cockpit procedural trainer that does not require an aerodynamic model but is accurate in systems modeling. Level 5 is for specific classes of aircraft and meet specific FTD design criteria for aerodynamic programming and systems modeling. Level 6 FTDs have high fidelity and are aircraft specific. Level 7 is for helicopter simulations. FTDs are currently in use in many university flight training programs, as well as at smaller flight schools (Beckman, 2000). The third type of training device is the Aviation Training Device (ATD), and PCATDs fall into this category. These devices typically consist of an aircraft control console that provides the flight controls necessary for performing flight maneuvers, a high performance desktop computer, and PC monitors as visual displays. There are two classifications of ATD: basic (BATD) and advanced (AATD). The AATD is more representative of specific aircraft types in terms of avionics displays, cockpit design that replicates the aircraft, and performance of the aircraft in terms of pitch, bank, and yaw (Beckman, 2000; FAA, 2018).

Advantages

Flight simulation training is widely used and provides several advantages for aviation training. These advantages include, (a) significantly reducing training costs compared to training in real airplanes, (b) providing a safe environment to practice potentially dangerous procedures, such as an engine failure or hard landing, (c) being eco-friendly by way of conservation of resources and reduced carbon footprint, (d) not being influenced by weather conditions, (e) allowing rapid and multiple repetitions of events, such as instrument approaches and landings, and (f) providing a research platform through easy-to-set-up scenarios for laboratory testing(Jorna, 1993; Williges et al., 2001).

Disadvantages

However, using flight simulation for training does have some disadvantages including (a) simulator sickness, a syndrome similar to motion sickness, that is often experienced during simulations, (b) providing poor or no motion cueing, (c) inducing adaptation and compensatory skills (d) having a complex system architecture requiring maintenance and troubleshooting, and (e) high costs associated with advanced simulators (Myers et al., 2018).

Simulation based training, Simulation fidelity, and Training Effectiveness

Today, simulations are widely used for training, evaluation, and analysis purposes (Thompson et al., 2008). As a training device, simulation provides trainees a safe environment to have hands-on practice for learning objectives. Compared to conventional lecture based training, simulation based training is using a constructivist approach, which means it uses active learning by creating meaning from experience (Ertmer & Newby, 2013). For simulation based training to be successful, trainees must be able to apply learned knowledge, skills, and abilities gained from simulation to real-world situations, which involves the transfer of training (Liu et al., 2008).

The FAA defines the transfer of training as the "ability to apply knowledge or procedures learned in one context to new contexts" (FAA, 2008). The transfer of training can be positive or negative. Positive transfer occurs when past learning helps the student with new learning. Negative transfer occurs when past learning interferes with and makes learning new skills more difficult (Sousa, 2017).

Transfer of training is often evaluated by objective performance measures or subjective judgments (Spector, 2003). Objective measures are data that objectively reflect the trainee's performance level, for examples, the number of errors, time to complete a task, and reaction time. Subjective measures are those that can include more potential variance or bias, such as ratings given by an expert or instructor on a trainee's performance (Liu et al., 2008). Several other calculations can also be done to assess transfer of learning, including percent transfer (the saving of time or trials in an aircraft by using a flight simulator); transfer effectiveness ratio (measures the efficiency of the simulation); first shot performance (how much training will be retained on first transference to the real situation); training retained (how much training is retained on the first posttransfer trial from the simulator compared with that gained from the real world). The selection of performance measures depends largely on the training tasks and experimental design (Liu et al., 2008).

Simulation fidelity is a key concept in simulation design; it refers to the degree of realism of simulation compared to the real activities (Myers et al., 2018). The Fidelity of a simulation can be broken down further into physical fidelity, cognitive fidelity, and functional fidelity. For flight simulation, physical fidelity refers to the level to which the simulation replicates actual physical aircraft flight characteristics, including motion, vision, and sound replication (Myers et al., 2018). For example, a flight simulation with high physical fidelity would be built using flight controls and instruments that are the exact copy of the real aircraft ones and they would be placed at the exact location where they were installed in the aircraft. Cognitive fidelity refers to the ability of the simulation training environment to replicate the cognitive skills required to perform flight activities. Specifically, psychological and perceptual factors such as situational awareness, anxiety, stress, and decision-making process will contribute to the cognitive fidelity (Lee, 2009; Taber, 2014). For example, a flight simulation with high cognitive fidelity should require the pilot to use the same attentional resources and produce similar psychological effects, such as stress and workload, compared to flying the real aircraft (Liu et al. 2008). Functional fidelity refers to how the simulator reacts to the tasks and commands being executed by the user compared to interaction with a real aircraft (Allen, 1986).

It is natural to assume that the higher the level of fidelity, the higher the degree of transfer of training will occur, however, there is considerable debate regarding the effectiveness of simulator fidelity on training transfer (Liu et al., 2008). Beckman (2000), Talleur et. al (2003), and McDermott (2005) proposed that there were no significant differences in training effectiveness between low fidelity PCATD and high fidelity FTDs for Instrument Flight Training. Dahlström (2008) discussed the fact that high fidelity simulation has not necessarily resulted in improved opportunities for learning coordinative and cognitive skills. Viden & Hall (2005) conducted a meta-analysis on the effect of simulator platform motion, and they concluded that a lack of motion caused trainees to be less successful in developing flight control strategies than those trainees who had practiced the skill with motion. Alessi (1988) proposed that the relationship between fidelity and learning is nonlinear; when the fidelity level is increased, the corresponding change in transfer of training depends largely on the trainee’s characteristics and ability to respond to this increase in fidelity. Noble (2002) also argued that as the learner skill level improves, low fidelity devices become less effective when one considers the cost to build them versus training efficiency. There is no easy answer to how simulation fidelity affects training effectiveness; it depends on many factors including the individual trainee's characteristics, the instructor, the training design, and the particular skills to be learned and transferred (Liu et al., 2008).

Kirkpatrick & Kirkpatrick (2006) proposed a four-level model of measuring training effectiveness. In this model, the first level is reaction, which reflects how the trainees perceive the effectiveness of training. It is important that the trainees enjoyed the training and felt that it was a valuable experience (Thomas, 2018). The second level is learning, which described the skills, knowledge, and principles understood by the trainees. This type of learning is often measured by knowledge and comprehension tests. The third level is behavior, which relates to how the trainees apply the information learned to real-world tasks. This type of measurement involves a longer period of evaluation (Kirkpatrick, 2006). The fourth level is result, which is measured by trainees’ achievement and implementation of the desired training goals over time to improve job performance and trainee morale.

Kraiger, Ford, and Salas (1993) developed another model of training evaluation. Their model classified training outcomes in categories: cognitive, skill-based, and affective learning outcomes. Cognitive outcomes are similar to the learning level in Kirkpatricks' model and include verbal knowledge, knowledge organization, and cognitive strategies. Verbal knowledge exists in different forms: declarative knowledge (information about what), procedural knowledge (information about how), and strategic knowledge. Verbal knowledge can be assessed through knowledge exams (Kraiger et al., 1993). Knowledge organization refers to how learners develop procedural knowledge and how that knowledge is organized in the trainees' minds. Knowledge organization helps trainees to develop mental models or cognitive maps. Evaluation of knowledge organization can involve a trainee writing down a mental model and comparing it to an expert mental model (Kraiger et al., 1993). Cognitive strategies refer to mental activities that facilitate knowledge acquisition and application. Measuring the usage of cognitive strategies could be done by giving an evaluation measure of an entire training program (Kraiger et al., 1993). Skill-based outcomes reflect the development of technical or motor skills, and those are comparable to the Behavior level in Kirkpatricks' model. This type of training outcome can be assessed by observing the trainee in a simulated or actual environment completing a trained task. Common measures for skill-based outcomes are, time to completion, counts of the number of errors made, or performance rating by subject experts (Kraiger et al., 1993). Affective outcomes are similar to the Reaction level in Kirkpatricks' model; they are the internal attitudes or motivations that can determine behavior or performance Training could cause a change in attitude, which leads trainees to pay more attention to learn and use the skills acquired in the future. This type of outcome could be measured by changes in trainees' self-efficacy perceptions and motivation by questionnaire (Kraiger et al., 1993).

The relationship between simulation and training effectiveness is complex. Thus, to investigate the effect of a simulation-based training method, it will be appropriate to use different types of measures to assess the different levels of training effectiveness.

Flight Simulation Research

Transfer of training research

Much research has been conducted to assess the transfer of learning in flight simulations (Combs, 2001; Gheorghiu, 2013; Zaa et al., 2015). Oritz (1993) conducted a study involving 60 participants with no previous flight experience at Andrews University. The participants were randomly divided into two even groups: an experimental group and a control group. The experimental group trained in a PCATD on a maneuver before performing the same maneuver in a real airplane. The PCATD consisted of a Zenith ZBO 3303 GQ Personal Computer, 14" VGA Color Monitor, CH Flightstick, Maxx rudder pedals, and ELITE simulation software. The PCATD was set to model the performance capabilities of a Cessna 150/152. The maneuver consisted of flying a squared pattern involving flying north, east, south, and west for 1.5 minutes each with right turns at the end of each leg, and a 450-degree turn to the right after the west leg, ending on a north heading. The experimental group was required to practice the entire maneuver with a flight instructor's verbal assistance until reaching the following limits: Altitude within ± 100 feet, Heading within ± 10 Degree, Bank angle within 10 Degrees. Then they were taken to the actual airplane to perform the maneuver to the same limits. The control group was taken directly to the airplane for the same training. Both groups used the same certified flight instructor, and their training time was recorded.

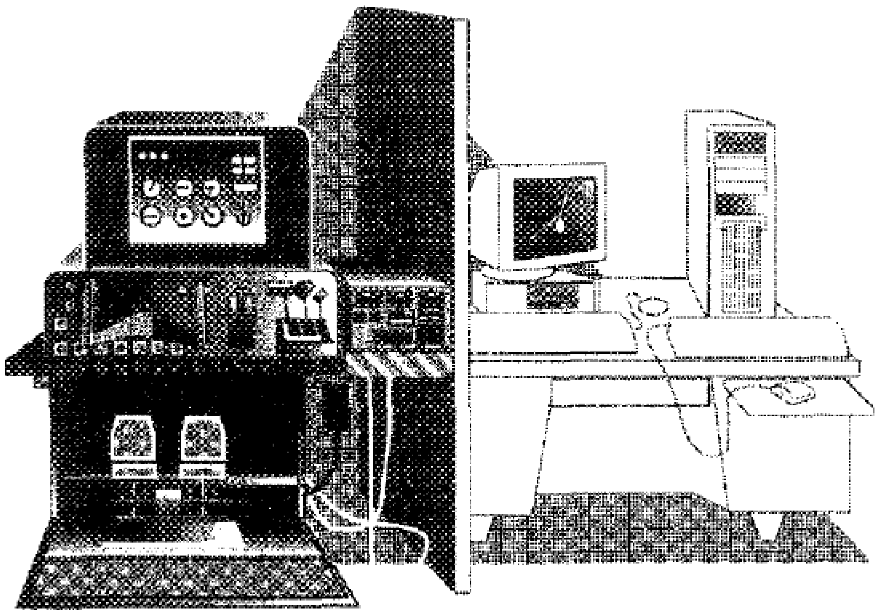
The transfer of learning was analyzed by the transfer effectiveness ratio (TER). TER is a measure of time savings in the aircraft as a function of time or trials in a training device. The

TER formula, developed by Stanley N. Roscoe (1980) is, TER = (Yc – Ye)/Xe. Yc is the average time (or trials) required by the control group to reach criterion in the actual aircraft, Ye is the average time (or trials) required by experimental group to reach criterion in the actual aircraft. Xe is the average time (or trials) required by the experimental group to reach criterion in the training device. The overall average time in minutes and seconds for the experimental group was 16:48 in PCATD and 12:23 in airplane, and for the control group 20:23 airplane. The study found the PCATD was as effective a teaching tool with a TER of 0.48 (Oritz, 1993).

The University of Illinois conducted a training experiment involving 107 student pilots (Taylor et al., 1999). Data collection started with the Aviation 130 class in fall 1994 and continued with the Aviation 130 and 140 classes in subsequent semesters through the Aviation 140 class in spring 1996. Participants were assigned randomly to the PCATD group and the airplane-control group with the constraint that male and female students were distributed evenly between the two groups. Fifty-four participants were assigned to the airplane-control group and 53 participants to the PCATD group as they entered the Aviation 130 class. Thirty-nine participants in the airplane-control group and 38 participants in the PCATD group continued into Aviation 140 class. The experimental PCATD consisted of an IBM-compatible Pentium 60 h4HZ Computer, an instructor-station map display, Precision Flight Controls, a 20" monitor permitted display of six standard instruments, MDM FS100 simulation software, and the PCATD was modified for Beechcraft Sundowner performance characteristics. The actual training airplanes were also Beechcraft Sundowners. Figure 1 showed the PCATD setup.

Figure1.

The computer to the right controls the FS- 100 simulation software. The computer to the left controls the instructor station (Taylor et al., 1999)



.

The students were scheduled for 45 hours of academic classes (ground school) and 15 flight lessons in each class. The objective of the class was to train the skills necessary to perform instrument flight rules (IFR) flight, including departure, en route, and arrival procedures (Taylor et al., 1999). For the PCATD group, all maneuvers and procedures were introduced and taught to proficiency standards in the PCATD prior to training in the airplane. For the control group, all maneuvers and procedures were introduced and taught to proficiency standards in the airplane. Instructors rated student performances on designated instrument tasks in both the PCATD and the aircraft for the PCATD group; for the control group, instructors rated student performances on those same instrument tasks only in the aircraft (Taylor et al., 1999). For performance assessment, instructors recorded trials to the criterion for specific tasks and instruction time to complete lessons. For both groups, to progress to a new flight lesson, the students were required to reach the standard proficiency in the airplane in the previous lesson.

Percent transfer and TERs were used for analyses in the study (Taylor et al., 1999). The percent transfer formula is $(Yc - Ye)/Yc \times 100$ = Percent transfer. For Aviation 130 class instrument tasks, Percent transfer values ranged from 11.2% to 33.3% and TERs ranged from 0.12 to 0.28. Only transfer values of 20.4% or higher and TER's of 0.25 or higher were associated with significantly different means for airplane trials in the two groups. For Aviation 140 class instrument tasks, Percent transfer values ranged from 13.2% to 28% and TERs ranged from -0.11 to 0.38. Transfer values of 14.6% or higher and TERs of .16 or higher were generally associated with significant differences between the two groups. For the Aviation 130 class, the PCATD group required a mean of 21 hours to complete the course, which was significantly shorter than the control group's 23.1 hours, $t(90)= 3.53$, $p< .001$ (two-tailed). For the Aviation 140 class, the PCATD group required a mean of 26.37 hours to complete the course, which was also significantly shorter than the control group's 28.18 hours. Overall, 3.9 hours were saved for the Aviation 130 and 140 combined, and the PCATD group used an average of 26.5 hours in total on the PCATD, making the cumulative TER time 0.15.

Overall, the results of the Taylor et al. (1999) study demonstrated that using a PCATD can positively transfer learning to the aircraft and showed that the values of percent transfer and TER changed significantly between maneuvers. They concluded that the PCATD was most effective at introducing maneuvers rather than practicing the maneuvers later in the course, as TER values decreased as instrument training progressed (Taylor et al., 1999).

Lintern, et al. (1997) conducted a study that used both a quasi -transfer and a transfer experiment design to evaluate the effects of Scene Detail and Visual Augmentation in Landing

Training on a simulator. The term quasi-transfer is used to characterize a study in which transfer testing is undertaken in a criterion configurated simulation or a high-fidelity simulation. For example, to design a transfer experiment for testing the relative training effectiveness of two simulation configurations. In a quasi-transfer experiment, testing in a high-fidelity simulation would be substituted for testing in the aircraft. Seventy-two male and 12 female flight students with no previous flight experience were tested in a transfer-of-training design. Fifty-four of the male students and 6 of the female students were included in a quasi-transfer design, which was nested within the larger transfer experiment (Lintern et al., 1997). The aircraft used were Beechcraft Sports and Sundowners. Two ILLInois Micro Aviation Computer (ILLIMAC) simulators were used in simulation training, one had real-time colored visual graphics system, and another had no visual graphics system.

The students were assigned randomly with equal distribution over the levels of landing training, (0, 24, 48, or 72 landing trials in the ILLIMAC with the visual display). In the simulation, “students were to make landing approaches, lined up with the runway centerline, on a 4' glideslope. The start points placed the students 0.5" below the 4.0" glideslope, 10,100 ft (3,078 m) from the runway threshold and 635 ft (190 m) above ground level” (Lintern, et al., 1997, *p*. 154). Students were required to maintain level flight until they captured the designated descent path. During the descent, they were to maintain an airspeed of 70 knots. The students with 24, 48, and 72 visual landing trials were assigned randomly with equal distribution over two levels of scene detail (moderate & low) and three levels of augmented guidance (off, constant, & adaptive.). In the low-detail scene, the objects and ground patterns were removed from the runway surrounding, and the moderate-detail scene provided some of the ground features associated with an airport. This visual guidance consisted of eight pairs of facing "F-poles" which defined the slope of the desired approach path to the runway aim point and an extended runway centerline. In the constant mode, the guidance was on throughout each trial. In the adaptive mode, the augmented guidance switched on only when students flew out of a predefined flight envelope and switched off again when they returned to a slightly tighter envelope (Lintern, et al., 1997).

The last 12 trials in simulation, which constituted a quasi-transfer test, were flown under the condition of moderate scene detail and no augmented guidance, which was the one that most closely approximated real flight conditions. Then, all students were taken to fly in an actual airplane to test the true transfer of training (Lintern, et al., 1997). For the quasi-transfer assessment, altitude deviations from the 4" descent path and lateral deviations from the runway lineup were collected at a 12-Hz sampling rate. Means and within-trial variances were calculated for each variable from 2,425 m to 606 m from the runway aim point. The dependent measures for the actual flight transfer phase were the number of student's attempted landings (SALs) prior to release for solo and the number of training sessions in which students attempted landings prior to release for solo (landing sessions). The study results showed no significant transfer effect of the number of training trials, however, compared to the data of the control group (who had received no additional simulator training) from the authors' previous similar study, the transfer effect existed. When training was conducted in the visual simulator, it required an average of 59.12 SALs compared with 73.44 SALs for the controls (Lintern, et al., 1997). The pattern of augmented guidance effects found in quasi-transfer for training with a moderate level of detail was repeated in transfer. The study discussed that correspondence between effects found in quasi-transfer and in transfer depended on the variables chosen to measure training. They also suggested using a higher-fidelity simulation in the future quasi-transfer test.

Atkins, Lansdowne, & Pfister (2002) conducted a study on control mechanisms in flight simulation with a quasi-transfer design. They examined the transfer of training between two differing flight control mechanisms (yoke & joystick) on a simulated visual landing approach. The dependent variable used in the study was vertical glideslope error (VGSE) in degrees of deviation per second while the simulated aircraft was in the air. Transfer effect was assessed between the acquisition and transfer phases one week apart in four conditions (8 participants per condition): Yoke to Yoke, Yoke-Joystick, Joystick-Joystick, and Joystick-Yoke. Overall results showed positive transfer effects between the transfer phase and acquisition phase performance. Although, a significant difference between control mechanisms was not found, interestingly, many participants reported that the joystick was an easier control mechanism to use than the yoke (Atkins et al., 2002).

O'Malley et al.'s study (2016) and Zaal et al.'s study (2015) both used a quasi-transfer approach to test the effects of motion on the transfer of training. Both studies used levels of motion conditions on a simulator as independent variables and used performance on the highest-fidelity motion condition simulation as the check of transfer of training. O'Malley et al. (2016) found that simulation training without disturbance motion cues might have produced better knowledge of the consequence of flight control movements, leading to increased sensitivity to the disturbance in the test phase. Zaal et al. (2015) found significant differences in pedal input reaction time and longitudinal deviation from the desired touchdown point between motion conditions. Both studies discussed the likelihood that dependent measures of the effectiveness of the transfer of training were influenced by participants' personal preferences and experiences.

Research Comparing Across Simulations

Beckman (2000) conducted a study to compare the effectiveness of a PCATD and an FTD for Instrument Flight Training. A two-group between-subject design was used for the experiment. The PCATD group (8 participants) received two sessions of holding pattern instruction in the Jeppesen FS-200 PCATD prior to demonstrating their skills in a TB-9 aircraft, and the FTD group (8 participants) received the same two instruction in the Fraca 141 FTD prior to testing in the aircraft. Altitude, heading, ability to track assigned radial, time inbound to the station, orientation during the holding pattern, and ability to become established in the hold were the parameters for the evaluation flight. A flight instructor rated all participant's performances based on the following criteria:

Students began with 100 points. Altitude off more than 100 ft, minus 1 point for each 3 seconds of deviation. Heading off more than 10 degrees while outbound, minus 1 point for each 3 seconds of deviation. More than 10 degrees from assigned radial while inbound, minus 1 point for every 3 seconds of deviation. Time inbound, minus 1 point for every five seconds deviation from one minute. Orientation, minus 5 points for each incorrect answer regarding orientation during holding pattern. Inability to become established in hold, minus 10 points for each unsuccessful circuit (Beckman, 2000, p.30).

The average score for the FTD group was 68.125, and the average score for the PCATD group was 70.5. T-test results found that the difference between the means of the two groups was not statistically significant, and the two training devices are equally effective in preparing a student for the task (Beckman, 2000).

Talleur et. al (2003) conducted a study involving 106 instrument-rated pilots examining the effectiveness of PCATD's for maintaining the Federal Aviation Administration's instrument currency requirement. The study employed a between-subject experiment design with four groups: aircraft, simulated in FTD, and PCATD trainer as well as a no training- control. Two FAA-approved Jeppesen FS-200 PCATD's, configured as Beechcraft Sundowner, two FAA-approved Frasca 141 FTD's, and two Beechcraft Sundowner aircraft (BE-C23) were used for the experiment.

All participants received an instrument proficiency check as the baseline (IPC 1) with a flight instructor in the actual aircraft to start the 6-month experimental period (Talleur et al., 2003). The IPC was a standardized test of the instrument pilot's skills in the aircraft. Then pilots in the aircraft, FTD, and PCATD groups received a practice training at the 2nd and 4th months of their participation, then did an IPC again in an airplane at the 6th month (IPCI 2). The control group received no training during the six month period prior to the second time IPC airplane flight. The effectiveness of the training device was assessed by comparing IPC 1 and IPC 2 pass/fail ratios and changes in maneuver performance rated by the flight instructor for the four pilot groups. The PCATD and FTD group in this study performed significantly better on IPC 2 than the control group, however, the Aircraft groups were not statistically different than the control group. The FTD group showed a slightly larger improvement than the PCATD group in pass/fail ratios between IPC 1 and IPC 2, but PCATD appears more effective when considering individual maneuver performance. Overall, this study provided evidence, reported later in this chapter, that the PCATD was as effective for instrument training as an FTD (Talleur et al., 2003).

McDermott (2005) conducted a study involving 63 instrument-rated pilots to compare the effectiveness of a PCATD and an FTD at improving pilot instrument proficiency on instrument landing system (ILS) approaches. Frasca FTD Model 141 and Precision Flight Controls PCATD Model PI 142 were used as training devices. The PCATD group received four training lessons with a PCATD and a final assessment on an FTD; the FTD group only used the FTD for the same training lessons and final assessment. Flight instructors were both the teacher of lessons and the rater for the assessment.

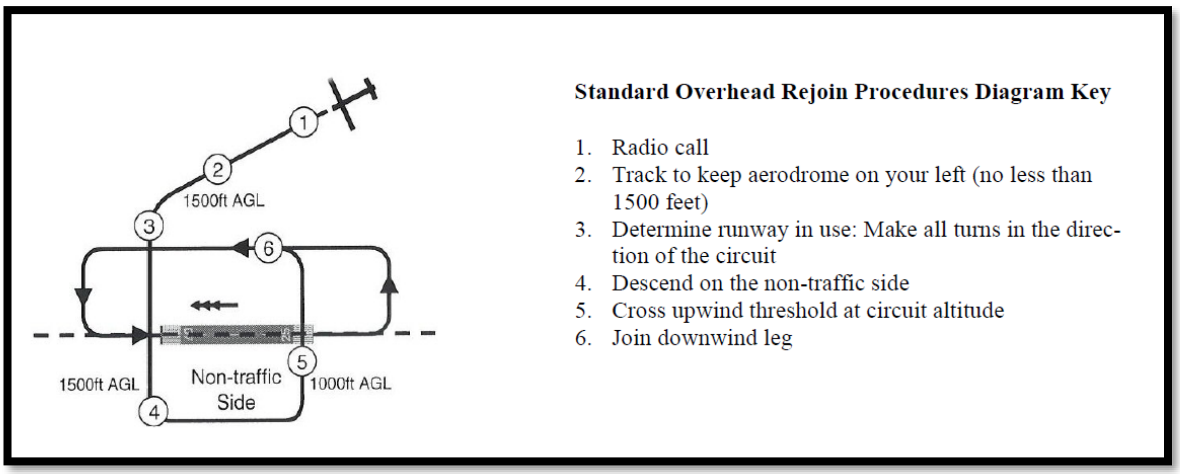
Measurements used for the analysis were the rating scores on participants' flight skills and experience in terms of total flight hours, actual instrument flight hours, flight simulation hours, recent flight hours, and numbers of instrument approach. Results indicated no statistically significant difference between using the PCATD and the FTD for maintaining an instrument pilot's ILS proficiency (McDermott, 2005). Additionally, a Post-Simulation Feedback survey was given to the participants to ask their opinion and experience with the training devices. Nearly half of participants (48%) thought that flight skills improved to a significant degree in the simulation lessons and 55% of participants indicated that their flight skills improved to a significant degree because their basic instrument flight skills of scanning, pacing, and instrument interpretation improved (McDermott, 2005).

Reweti et al. (2017) conducted a quasi-transfer study to compare training effectiveness between PCATD's and certified FTD's at improving pilot proficiency in the performance of a standard visual flight rules (VFR) traffic pattern. Ninety-three pilots were first randomly assigned to one of three groups (PCATD group, FTD group, and the control group). A pre-test was administered to each group; the participants completed a standard VFR rejoin procedure (showen in Figure 2) on the Frasca TruFlite Flight & Navigational Procedures Trainer, commonly referred to as the Frasca FTD. Then, the PCATD group received training on a PCATD running Microsoft Flight Simulator X, the FTD group received training on the same

Frasca FTD, and the control group received no training. The PCATD and Frasca FTD were configured as single-engine PA-28 Piper Warrior airplanes. After training, each group of participants was given a post-test, which was identical to the pre-test procedure using the Frasca FTD (Reweti et al., 2017).

Figure 2

Diagram of Standard Overhead Rejoin (Reweti et al., 2017).



The flight data were recorded and scored using the National Intercollegiate Flying Association (NIFA) Score Editor. The NIFA Score Editor is a program used to measure the performance of pilots as they attempt to fly an established flight pattern; the program can record the number of errors committed by participants across a number of flight variables including Pitch, Bank, Altitude, Indicated airspeed, Heading, Glide slope, Overhead rejoin pattern, and a Total variable (combined score of Pitch, Bank, Altitude, IAS, and Heading), and then give a score for the performance (Reweti et al., 2017). A high score (e.g., 20 penalty points per second) represents a high number of errors and a poor performance. A series of 3 x 2, mixed-model

ANOVA tests were used to explore if there were statistically significant differences between the pre-test score and the post-test score between three groups on eight performance variables. The results suggested that the use of both PCATD and FTD led to significant improvements in VFR task performance compared to a control group, and there was no significant difference in pre-test /post-test change scores across all of the eight variables between the FTD group and the PCATD group (Reweti et al., 2017).

Virtual Reality

Virtual reality (VR) refers to an experience, in which a user is surrounded by a computer-generated immersive virtual environment (IVE) that one can navigate and possibly interact with, resulting in real-time simulation (Brennesholtz, 2018; Oberhauser et al., 2018; Sacks et al., 2013). The origins of VR technology reach back to 1965 when Ivan Sutherland developed the first head-mounted display (HMD; Oberhauser et al., 2018). However, with limited computing, it was not possible to deliver a satisfying experience at a reasonable price until recent years (Robertson & Zelenko, 2014). As technologies advanced, today hardware and software have emerged into more compact and affordable VR solutions with high quality HMD, position tracking systems, versatile control input, and high computer processing and graphical power (Geršak et al., 2018).

VR experiences are typically delivered by using an HMD with 360° sound that could temporarily reduce or remove the user's perception of the real environment. Some advanced VR systems also include haptic displays (Bowman & McMahan, 2007). The VR systems generally track the motion of a user's head or hand-held controls, and the received data is used to determine the user's view, navigation, interaction with objects, and possible movement of the "avatar" in the IVE (Burdea & Coiffet, 2003; Brennesholtz, 2018). Generally, there are two types of VR, 3 degrees-of-freedom (3DOF) VR and 6 degrees-of-freedom (6DOF) VR (Brennesholtz, 2018). In 3DOF VR, the viewpoint is fixed by the content and the user can only change the viewpoint direction. The system responds to the angular motion of the user's head in roll, pitch, and yaw; the term 360° video is sometimes used interchangeably with 3DOF VR. However, 3DOF VR typically allows the user to interact with his visual environment while in 360° video, the user can only look at the preprogrammed video (Brennesholtz, 2018). 3DOF typically can be viewed on a smartphone-based HMD (e.g. Samsung Gear VR) or a dedicated HMD (e.g. the OculusVR). In 6DOF VR, the user can control not only the three angular dimensions of the viewpoint direction but can control the three spatial dimensions (x, y and z axis) of the viewpoint position itself, giving the 6 degrees of freedom, in other words, 6 DOF VR gives the user more freedom to move the "avatar" in the virtual environment. (Brennesholtz, 2018). 6DOF VR is typically viewed on a dedicated HMD cable or wirelessly connected to a PC with play area trackers (e.g. the HTC Vive).

Immersion refers to the objective level of sensory fidelity a VR system provides. The level of immersion depends only on the system's rendering software and display technology and is one main factor that influences a user's subjective psychological response or the feelings of "presence" while using a VR system (Bowman & McMahan, 2007). In a fully immersive VR system, users' responses in the virtual environment are similar to their responses in a real environment, and users can receive many depth cues which can lead to greater spatial understanding.

Use of VR in Training and Learning

One of the most popular fields of VR application is training. VR is useful for single person interaction with highly detailed tasks or settings, and the virtual environments can effectively simulate various conditions of work and life while, successfully supporting learning processes (Sanchez-Vives & Slater, 2005; Norris et al., 2019). High fidelity sensory stimuli present in VR simulations play a role in their success (Bowman & McMahan, 2007). Parong and Mayer (2018) argued that using immersive VR for teaching is grounded in interest theory and self-efficacy theory; immersive VR would generate learner's situational interest which leads them to pay closer attention to the learning content. VR may also increase the student's self-efficacy by providing appropriate feedback from virtual interaction, which enhances a learner's motivation for study. Perhaps, VR flight simulation is a valuable tool to improve the Reaction level of training effectiveness in Kirkpatricks' model or Affective outcomes in Kraiger, Ford, and Salas' model.

There are many studies conducted to evaluate the effectiveness of training methods using VR technology. Sacks et al. (2013) tested the hypothesis that safety training in VR would be feasible and effective for construction safety training. Sixty-six participants were provided training in construction safety and their safety knowledge was tested prior to the training, immediately afterward, and one month later. Participants were divided into two groups, a traditional training group with classroom instruction using slides shows and a VR training group. The training content was collated from safety codes and from the standard construction supervisors' safety courses run by the Israel Institute for Occupational Safety and Hygiene. For the VR group, 21 training scenarios were created in a virtual construction site. Participants also were asked to fill out an experience questionnaire immediately after the second safety test post-training.

T-test results between the pre and post-training safety knowledge test scores showed significant improvement; the results demonstrate effectiveness in immediate learning of hazard identification and prevention skills for both the VR and the traditional training groups (Sacks et al., 2013). The margin of improvement of the test scores for each type of training was compared. VR was significantly better than traditional training for hazard identification for reinforced concrete works, stone cladding works, and learning prevention knowledge for the cast in-situ concrete works. The short-term effectiveness of the training was investigated by comparing scores for the safety tests administered before training and one month after training. With-in group comparison showed both training methods to be effective in scores on risk for reinforced concrete works identification for reinforced concrete works and stone cladding works and in scores for prevention for general site safety and stone cladding works. However, since only 23 participants returned to complete the third safety knowledge test, between-group comparisons did not have enough power to show an advantage for virtual reality training. The learning experience questionnaire results showed a significant advantage for VR over traditional training. Additionally, the researcher who observed the traditional training sessions noted that trainees tended to lose concentration after about 40 minutes. In contrast, the virtual reality trainees were observed to maintain full focus for the whole one and half hour training session. Overall, the researcher concluded that instruction using VR was more effective than safety training with traditional classroom presentations (Sacks et al., 2013).

University of California, Santa Barbara conducted a study to compare the instructional effectiveness of immersive VR and a desktop slideshow as media for teaching scientific knowledge, as well as, to examine the efficacy of adding a generative learning strategy to a VR lesson (Parong & Mayer, 2018).

In their first experiment, college students learned a biology lesson about how cells in the human bloodstream work either in immersive VR or use a self-directed PowerPoint slideshow on a desktop computer. There were 27 students in the VR group and 28 students in the slideshow group.

The VR lesson was presented to students using an interactive biology simulation software called The Body VR: Journey Inside a Cell, on a Dell Alienware computer with an HTC Vive VR system that included a head-mounted display and two wireless hand controllers (Parong & Mayer, 2018). The VR simulation contained narration and immersive animations of the circulatory system and parts of cells. In the simulation, the user traveled through an artery on a moving platform while a narrator explained the purpose of the cells within it. The slideshow lesson was adapted from The Body VR: Journey Inside a Cell; the same verbal content from the simulation was transcribed and printed in a slideshow format with corresponding screenshots from the simulation. For both groups, the students completed a prequestionnaire about their knowledge of the human body, a postquestionnaire about their experiences with the lesson, and a post-training knowledge test on the material they viewed during the lesson.

Experiment 1 results showed that the slideshow group scored significantly better than the VR group on the posttest factual questions, but not on the conceptual questions. However, the VR group rated the learning experience significantly higher in the postquestionnaire, on enjoyment, engagement, and motivation. They were more excited focused and less bored during the lesson than the slideshow group (Parong & Mayer, 2018).

In their second experiment, there was a VR group (29 participants) and a VR Plus group (28 participants). They used a similar procedure to the Experiment 1 VR group and received the same immersive VR to learn the lesson. However, for the VR Plus group, the lesson was divided into 6 segments, and the participants were asked to write a summary of the segment they just viewed after each segment (Parong & Mayer, 2018).

Experiment 2 results showed that the VR Plus group scored significantly better than the VR group on both factual questions and conceptual questions in the posttest. There was no significant difference in ratings on learning experience between the groups. Parong & Mayer (2018) argued the effectiveness of VR could be increased by prompting students to use generative learning strategies, such as summaries, without diminishing the learner's motivation, interest, engagement, and affect while using the new technology.

Vankipuram et al. (2010) developed a VR simulation for orthopedic bone drilling training and conducted a validation study. The VR simulation they used in the experiments has 3DOF and a haptic device interfaced with a surgical drill. Before they conducted the study, board-certified orthopedic surgeons were invited to do initial testing (Vankipuram et al., 2010). In the first experiment, six expert surgeons, 11 residents, and 6 novices were included. Participants were first asked to do two drilling tasks in the VR simulation with extensive guidance from researchers as the habituation practice. Then, they were asked to do four drill tasks that were to drill through a bone at a predefined marked area of the bone referred to as the target spot. The time taken to complete the tasks and the number of tissue contact errors made were measured as the objective performance metrics. The result showed a learning curve that time taken to complete trials decreased for all participants over the 4 trials. Expert surgeons made tissue contact errors at the beginning, converging to no errors at the end of four trials. Novices showed a constant high error rate while residents showed some improvement. After the experiment, the expert surgeons reported that at the beginning they were unfamiliarity with the VR simulation interface and they were cautious in not trying to speedily rush through the process and make errors. The novice and resident groups focused on time elapsed as a major variable of interest.

In the second experiment, 10 novice participants with no experience in orthopedic surgical bone drilling were divided into two groups (Vankipuram et al., 2010). The control group was directly put to the test of drilling through bone, and the experiment group was trained on the VR simulation prior to the test. All participants in the experiment group were required to reach zero tissue contact errors. In the testing stage, two groups were asked to perform drilling tasks on two identical bones. Three senior orthopedic surgeons were recruited to rate the participant's performance on a 10- point Likert scale. Positional error (the distance between the center of the target and the center of the drilled hole) was also measured for between-group comparisons. As result, the experiment group performed significantly better than the control group on both the objective and subjective measures. Researchers concluded that VR simulation is a valid training tool for orthopedic basic skills; the training in VR was transferrable to real environments and enhanced the understanding of the surgical procedures (Vankipuram et al., 2010).

Verdaasdonk, Dankelman, Lange, & Stassen (2008) conducted a study to assess the transfer validity of laparoscopic knot-tying training on a VR simulator to a realistic environment. Participants were recruited from first- and second-year surgical trainees who enrolled in a laparoscopic basic skills course. Each participant had as much training as basic skill training on the VR simulator until achieving the defined performance level, and then they were divided into two groups. In the experimental group (n=9), the participants received extra exercises on the knot-tying module on the laparoscopic VR simulator with video instruction on how to tie the double surgical knot on the simulator. The experimental group was required to tie a double surgical knot successfully on the laparoscopic VR simulator at least ten times before the testing. In the control group (n = 10), the participants received no further on hand VR training and only viewed three consecutive video demonstrations of the VR knot-tying procedure on the simulator before the testing. In the testing, both groups were asked to tie a double laparoscopic knot on an anesthetized porcine model (Verdaasdonk et al., 2008), the entire process was recorded digitally and coded for each individual. Objective assessments include Total time taken to tie the knot, numbers of errors, numbers of attempted loops, numbers of the needle tip touched the tissue, etc. Additionally, subjective assessments were made independently by two expert laparoscopic surgeons using a global rating scale. Results suggested that the experiment group was significantly faster with a significantly lower number of errors than the control group. The researcher concluded that the VR simulation was a useful tool to train laparoscopic knot-tying (Verdaasdonk et al., 2008).

VR Flight Simulation Related Research

With the advent of VR technology, the VR flight simulation was envisioned (Oberhauser et al., 2018). Some demonstrators of VR flight simulation include a virtual cockpit for a distributed interactive simulation created by the Air Force Institute of Technology (Mccarty et al., 1994.), a reconfigurable virtual research cockpit utilizes an HMD created by the University of California, Davis (Joyce & Robinson, 2015), a VR helicopter simulation created by the Middle East Technical University (Yavrucuk et al., 2011), and a VR flight simulator created by Oberhauser and Dreyer (2017). The main purpose to develop those VR flight simulations was to simulate the environment of real flight while reducing cost. Figure 3 and Figure 4 showed the reconfigurable virtual research cockpit and the VR helicopter simulation.

Figure 3

The Rapidly Reconfigurable Research Cockpit (Joyce & Robinson, 2015)

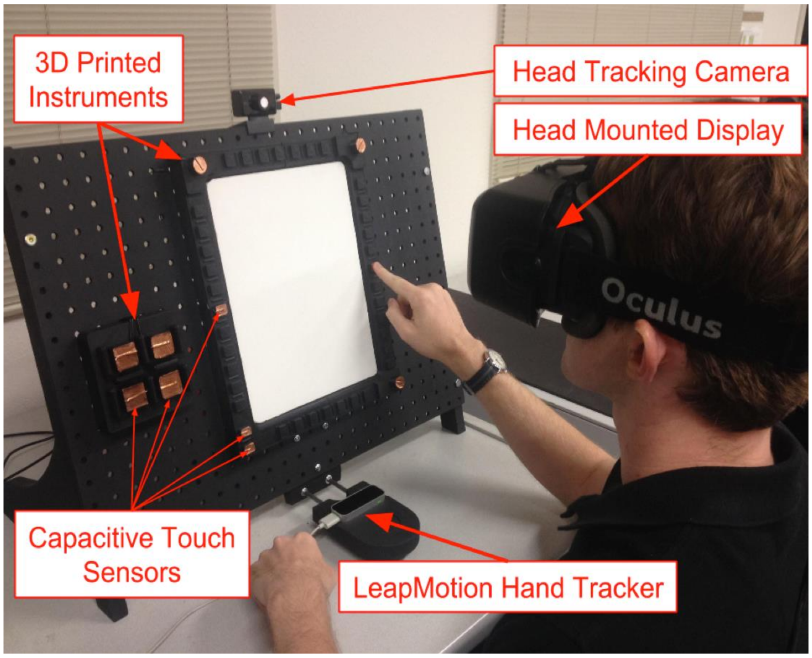
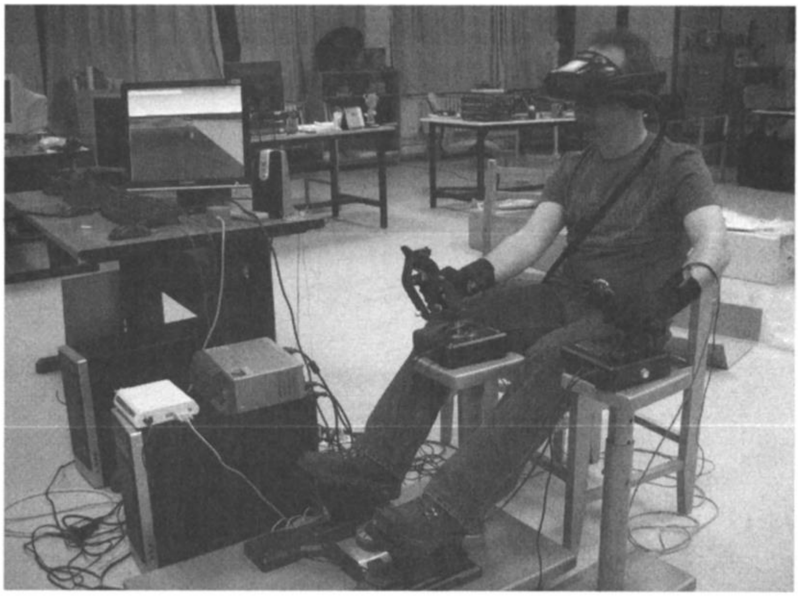


Figure 4

Virtual Reality Simulator (Yavrucuk et al., 2011).

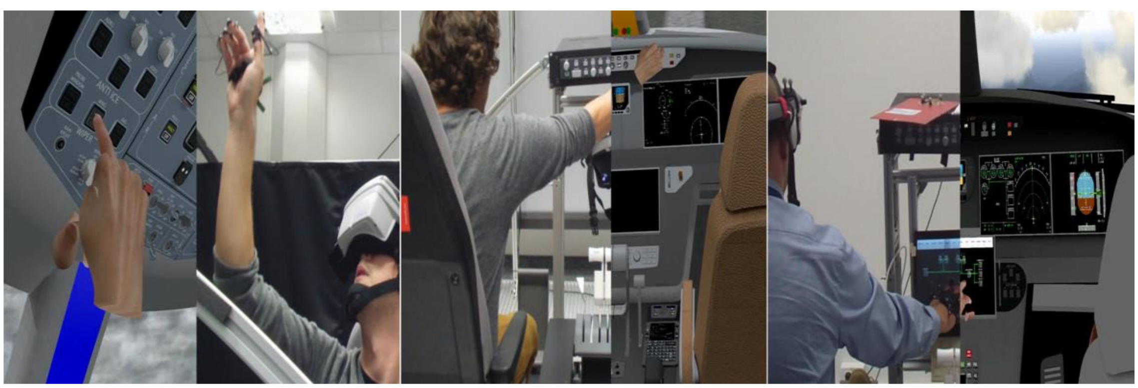


However, some early published research that involved using VR flight simulation was conducted in the field of psychological disorder therapy. Mu¨hlberger et al. (2001) conducted a study that examined the effects of repeated exposure of flight phobics to VR flight simulation. The result showed that VR flight exposure had a positive effect and greater fear reduction in flight to relaxation training. Rothbaum et al. (2006) compared the exposure therapy effect using VR flight simulation and using real airplanes on fear of flying. They concluded that VR simulation was essentially equivalent to exposure to the real airplane in terms of exposure therapy effect and had a significant positive effect than the no treatment control group.

Oberhauser and Dreyer (2017) developed a VR flight simulation for human factors engineering using a mixed mock-up and conducted a series of experiments involving 19 experienced commercial airline pilots and 12 non-pilots to evaluate the fidelity and usability of the VR simulation compared to a full flight simulator. Figure 5 showed the mixed mock-up for the VR flight simulation.

Figure 5

*Mixed mock-up and full virtual interact*ion (Oberhauser & Dreyer, 2017)



The experiment supervisor took the role of the co-pilot to ensure conformity with the flight tasks. Participants' heart rate and eye tracking heat maps were also collected during the experiments. The results showed that users' overall operational behavior in VR is comparable to the full flight simulator environment; the VR flight had a sufficient level of simulation fidelity and a sufficient level of usability to fulfill tasks like flying the aircraft and pressing numerous virtual and non-virtual buttons, as well as the interaction with a touch screen prototype, had to be completed in a time-critical scenario (Oberhauser & Dreyer, 2017). The VR flight simulation could be a valuable tool to gather reliable information on human factors aspects of the interaction in flights. The researchers also discussed some limitations of the VR simulation, for example, user's movements were slower and took more time to complete the task due to degraded ability to aim for the virtual non-haptic buttons, some users reported a low level of wearing comfort of the VR equipment, and users may experience simulation sickness in VR flight (Oberhauser & Dreyer, 2017).

Oberhauser et al. (2018) conducted a study that compared the pilot performance in a VR flight simulation and in a conventional flight simulation environment. The cockpit of a hardware simulator was remodeled and integrated into the VR simulation as shown in Figure 6. To use the VR simulation, a user needs to wear an HMD and tracking targets attached to the hands. In this configuration, the user can interact with the control element in the VR environment, and simultaneously touches the control element in the real hardware, which leads to the respective haptic sensation (Oberhauser et al., 2018).

Figure 6

A user immersed in the Virtual Reality Flight Simulator (Oberhauser et al., 2018).

The study used a within-subject design; the same 28 pilots with a mean age of 42.5 years and average total flight time of 2,485 hours performed the same flight tasks once in the VR simulation and once in a conventional FTD with the projected outside visual (Oberhauser et al., 2018). Before the experiment, the participant flew a left-hand pattern for familiarization with each simulation environment. In the experiment, the flight task starts with a short taxiing phase from the parking position, then takes off and flies a left-hand traffic pattern at an altitude of 2,000 ft. During the scenario, the participants receive pre-recorded audio instructions to interact with cockpit elements. After a touch-and-go and a second left-hand pattern, the participants were asked to land with a full stop.

Dependent measures used in the study included the Movement Time (the timespan of the hand traveling toward the control element, after the pilot has received the audio command), the Heading deviation, the Altitude deviation, the Lateral touchdown deviation, Runway heading alignment error, and Final approach cross-track error compared to the ideal traffic pattern, the

Pilots' workload using the NASA-Task Load Index, and the Simulator Sickness Questionnaire (SSQ; Oberhauser et al., 2018).

Oberhauser et al. (2018) found that the movement time in the VR simulation was significantly longer than in the conventional flight simulation for most interactions with the flight control elements and the deviation in-flight performance was significantly larger in VR than in the conventional flight simulation. The pilots' mental, physical, temporal demand, effort, and frustration on the NASA-Task Load Index were significantly higher in the VR than in the conventional flight simulation. In addition, participants' self-rating of performance was also significantly lower in VR. Participants' symptoms of simulator sickness were significantly stronger in VR. The researchers discussed that the difference in flight performance might be influenced by various confounding factors, for examples, the participants were unfamiliar with the VR interface, the inaccuracies in the virtual hand model, the VR display had more than 50 ms latency to the control input, and the HMD only offers a limited FOV (Oberhauser et al., 2018). However, the degradation of in-flight performance is not critical to safely conducting the given flight task in the VR environment. Even though the current VR flight simulation had several disadvantages compared to the conventional flight simulation, with the advances in the field of VR technology, the further development VR flight simulation still has the potential to be a valuable tool for training and research purposes, and how such VR flight simulation can be certified as flight training devices is subject to future research. (Oberhauser et al. 2018).

Summary

The flight simulations available for pilots today have many improvements in terms of fidelity, usability, convenience, cost efficiency, and variety by comparison to that of the past several decades (Page, 2000). With no doubt, with the advent of technology, the future development of flight simulations could present pilots with an even greater fidelity flight experience for training and practice.

The literature review includes many studies that compared the training effectiveness of different types of flight simulations. Overall, researchers found flight simulation had a positive effect for training purposes, especially for concept and procedure learning (Taylor et al., 1999). At the same time, many researchers found the effects of VR in medical, high-hazard work, military training, or generally science learning being positive (Greunke & Sadagic, 2016). However, the effect of VR flight simulation training and its comparison to traditional PCATD has not yet been studied and fully understood.

Lessons learned from existing research added valuable information in the development of the current study examining the effectiveness of VR flight simulation training. General methods to test and measure training effectiveness include knowledge tests, calculated transfer effectiveness ratio, flight performance on a real airplane or on a high-fidelity flight simulator (quasi-transfer study design) rated by flight instructors or measured by the flight parameters (e.g. vertical airspeed, altitude, deviation from the ideal flight path). The user experience of the flight simulation can be gathered using a post-training survey. For the current study, it is hypothesized:

H_{01}: All three groups (VR, Desktop, & Control) have no differences among the maneuver performance on the FTD

H_{02}: All three groups have no difference among the scores on knowledge tests (cognitive learning outcomes).

$H_{03:}$ User experience among all three groups has no difference.

$H_{04:}$ The trainees' motivation to use VR flight simulation for training and learning purposes is not different than motivation for use of conventional Desktop Flight Simulation training techniques or training techniques used in the control group.

$H_{05:}$ User experience is not related to the maneuver performance on the FTD.

H_{06}: The trainees' self-efficacy pertaining to the trained flight maneuver will have no change after practice in all three groups.

H_{07}: There is no difference in post-training self-efficacy on the selected flight maneuver among three groups.

$H_{08:}$ The trainees' self-efficacy pertaining to the trained flight maneuver is not related to the maneuver performance on the FTD in all three groups.

Method

Design

The current project employed a quasi-transfer of training study design. Quasi-transfer of training study design differs from the traditional transfer of training studies in that a high fidelity flight simulation rather than an aircraft is used to test training tasks. Quasi-transfer of training has been used successfully in a number of flight simulation experiments (Taylor, Lintern, & Koonce, 1993). The study has one independent variable with three levels and multiple dependent variables. This project used both quantitative and qualitative measurement approaches.

Participants

Advertisements for participants were distributed to Daytona Beach local flight schools and Embry-Riddle Aeronautical University's SONA and ETA systems. Forty-eight participants were recruited and were paid 30 dollars for their participation. To join this study, the participants were required to have Private Pilot License (PPL) or be in the middle of PPL training (passed SOLO) but haven't had commercial pilot training yet.

Table 1 below provides the demographic information about participants. Due to scheduling constraints, the number of participants was not equal in the three groups. The ANOVA tests showed that there was no significant difference among all three groups in terms of Logged Flight Hours, FTD Hours, or VR Flight Hours.

Table 1

Demographic Information for VR, Desktop, and Control Groups

	VR Group	Desktop Group	Control Group
Number of Participants	18	16	14
Participant Age	*M*= 19.33, *SD*= 1.37	*M*= 20.31, *SD*=2.44	*M*=22.14, *SD*=5.71
Participant Gender	Male= 14, Female=4	Male=14, Female=2	Male=11, Female= 3
Logged Flight Hours	*M*= 111.01,*SD*=58.44	*M*=100.41, *SD*=32.75	*M*=118.96, *SD*=59.22
FTD Hours	*M*= 20.49, *SD*= 27.01	*M*= 20.54, *SD*= 14.67	*M*= 20.40, *SD*= 15.21
VR Fight Hours	*M*= .06, *SD*=.24	*M*= 2.61, *SD* = 7.35	*M*= .43, *SD*= .94

Apparatus and materials.

The training task that was chosen for this study is a flight maneuver from commercial pilot curriculums, the Chandelle. Figure 7 demonstrates the flight pattern of a Chandelle. The standard procedures of a Chandelle are (Embry-Riddle Aeronautical University, 2018):

1. Select an altitude that will allow the maneuver to be completed no lower than 1500 feet above ground level.

2. Perform CLEARING TURNS and make a position report.

3. Adjust the pitch and power to maintain altitude and 105 KIAS (approx. 2350 RPM). Re-trim as necessary.

4. Select a prominent visual reference point directly off from the wing tip (left or right) and out towards the horizon.

5. Initiate a roll into a 30° bank in the direction of the reference point.

6. After the bank is established, initiate a climbing turn by smoothly applying elevator backpressure to increase the pitch attitude, and apply full power.

7. While maintaining a 30° bank, continue increasing the pitch attitude at a constant rate so as to attain the highest pitch attitude (approx. 13-15°) at the 90° point (reference point) in the turn.

8. At the 90°point in the turn, maintain the pitch attitude by continuing to increase backpressure (due to decreasing airspeed) and initiate a slow rate of rollout so as to arrive at the 180° point with the wings level (reference point off from the opposite wing) and at minimum controllable airspeed (stall warning).

9. To recover, maintain the heading while decreasing the pitch attitude, allowing the airplane to accelerate while maintaining the last altitude attained.

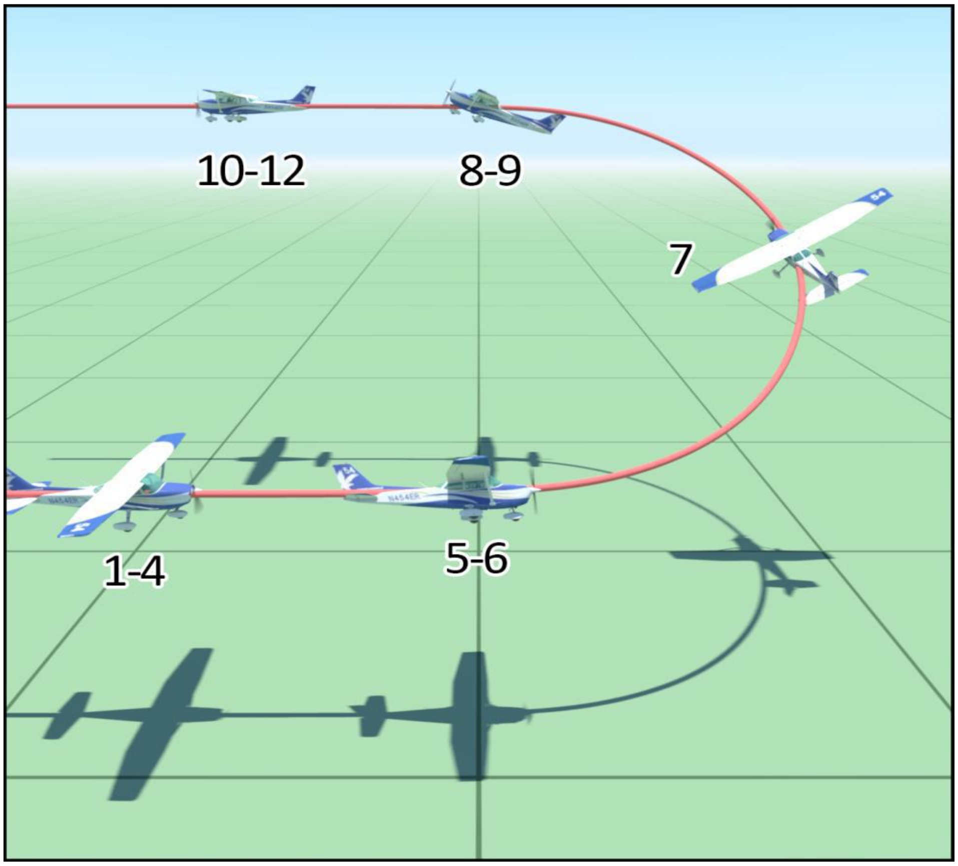
10. Set cruise power.

11. Re-trim as necessary.

12. Complete the CRUISE checklist.

Figure 7.

Chandelle (Embry-Riddle Aeronautical University, 2018).



Appendix A includes the written instructions of the Chandelle maneuver that was given to all three groups before their training activities (FAA, 2016). The Chandelle is a climbing U-turn, and the turn can be divided into two phases. Phase 1 starts when the plane initiates a roll into a 30° bank in the direction of the reference point and ends at the 90°point in the turn. For Phase 1, the pilot should maintain a 30° bank and continue to increase the pitch attitude at a constant rate. Phase 2 starts when the plane is at the 90°point of the turn and ends when the plane arrives at the 180° point with the wings level. For Phase 2, the pilot should maintain the pitch attitude and a slow rate of bank reduction.

There are three main reasons to choose the Chandelle for this study. First, the Chandelle is a relatively short maneuver and can be performed in flight simulations. Second, the Chandelle is a flight maneuver from commercial pilot curriculums. Thus, most private rating pilots will likely have little experience with such a maneuver. The third reason to select the Chandelle is that, although the Chandelle is a commercial maneuver, most PPL student pilots should already have the basic flight skills to perform the Chandelle.

The video instruction of the flight maneuver that was used for all three groups was retrieved from a YouTube video made by the University of North Dakota (2012).

The PCATD used for the Desktop group was a windows-based personal computer running X-Plane 11 with flight control accessories and three monitors. X-Plane 11 is a commercially available flight simulation application that supports both conventional PC interfaces and VR. The VR flight simulation used for the VR group was the same windows-based personal computer running X-Plane 11 in VR setting with the same flight control accessories and an HTC Vive Pro VR kit. Both flight simulations were configured to represent a Cessna 172 aircraft. A model airplane was used to assist the control group that orally demonstrated the Chandelle maneuver. Figure 8 was shown a participant using the VR flight simulation for practice.

Figure 8

A participant using the VR flight simulation

The testing phase of the current project took place at Embry-Riddle Aeronautical University's Advanced Flight Simulation Center. The FTD that was used for the post-training maneuver test was a Frasca C172S FTD with a data recorder. The FTD was also configured as a Cessna 172 aircraft.

Measures

The measures were a demographic questionnaire, user experience surveys, a knowledge test, a self-efficacy questionnaire, and flight performance data.

The demographic questionnaire was completed before the first training activity. The User experience Survey was completed after the training activity. Recorded flight parameter data on the FTD maneuver test was collected l. The post-training Chandelle maneuver knowledge tests were completed after the FTD maneuver test.

The demographic questionnaire used in this study (Appendix B) included questions about participants, such as age, gender, pilot rating, information about their flight experience specifically, total flight hours, years of flight experience, hours in simulation hours, as well as experience with VR applications or VR flight simulation and their willingness to use VR and PC based flight simulation for learning.

The user experience surveys (Appendix C) used for the study included the User Experience Questionnaire (UEQ; Laugwitz et al., 2008) and customized questions about users' preferences and willingness to use flight simulations for future training. The UEQ is a commonly-used user experience assessment tool for interactive products and was used for a number of VR studies (Anton et al., 2018; Su, Chen et al., 2019; Somrak et al., 2019). The Cronbach's alpha coefficients of UEQ subscales are .89 for Attractiveness, .82 for Perspicuity, .73 for Efficiency, .65 for Dependability, .76 for Stimulation, and .83 for Novelty (Laugwitz et al., 2008). The Cronbach's alpha coefficient of current project sample was .90 for Attractiveness, .77 for Perspicuity, .72 for Efficiency, .73 for Dependability, .79 for Stimulation, and .86 for Novelty.

The post-training Chandelle maneuver knowledge test (Appendix D) consists of written questions about the trained maneuver. The purpose of using the knowledge test is to assess participants' cognitive outcomes resulting from training.

The self-efficacy questionnaire (Appendix E) was designed by adapting items from the General Self-Efficacy scale (Schwarzer & Jerusalem, 1995) for the training activity. The questionnaire has 3 items, and each item would be rated by participants on a 7-point scale (1 as not at all, 7 as very). The purpose of using the self-efficacy questionnaire was to assess participants' affective outcomes resulting from training.

The flight parameters: bank, roll (bank) rate, pitch, pitch rate, heading, turn rate, and altitude were recorded every 0.2 seconds as a measure of training performance. The purpose of using flight parameters was to assess participants' skill-based outcomes of training. Flight performance on the FTD was evaluated by the deviation of those flight parameters during the maneuver. For example, the pitch should be maintained during the second half of the Chandelle maneuver for optimal performance. Thus, a smaller deviation indicated better performance.

Procedure

Before the experimental session, the participants completed the demographic questionnaire online, which included basic demographic information, pilot license rating, total flight hours, experience with flight simulation, and experience with VR applications. Upon arrival at the experimental site, participants were greeted and asked to review and sign the Informed Consent. Because the experiment was conducted during these uncertain times, there was a risk of contracting COVID-19. Both experimenter and participants were required to wash their hands before beginning the experiment and touch nothing between the bathroom and the research area. After each session, the experimenter used a disinfectant wipe to wipe all surfaces of the equipment used that were touched by the participant, as well as the pens, clipboard, and any other items.

Then participants were assigned to use one of three techniques (3 experimental groups) to complete the training session. The training technique was the main independent variable (IV) in this study. The three levels of the IV are Control, Desktop, and VR.

All three groups started training with reading a written explanation of the Chandelle maneuver (written learning (10 minute). Then, all three groups watched a video instruction of Chandelle maneuver (15 minute).

The first group was referred to as the Control group. After viewing the video instruction, they received no hands-on simulation training. Instead, they were asked to use a model airplane to demonstrate the Chandelle maneuver as practice orally for 2 trials (20 minutes). After each trial, the participant could go back to review the written instruction.

The second group was referred to as the Desktop group. After viewing the video instruction, they received a 5-minute initial practice training with the PCATD (desktop computer flight simulation) to become familiar with the flight simulation interface and control. Then, they were asked to practice the Chandelle maneuver for four trials (20 minutes). After each trial, the participant could go back to review the written instruction.

The third group was referred to as the VR group. Similar to the Desktop group, after viewing the video instruction, they received a 5- minute initial practice training with the VR-configured PCATD (desktop computer VR flight simulation) to familiarize themselves with the flight simulation interface and control. Then, they practiced the Chandelle maneuver for four trails (20 minutes). After each trial, the participant could go back to review the written instruction.

Additionally, all three groups were asked to complete a Self-Efficacy questionnaire right after reading the written instruction of the Chandelle maneuver and once again after the training practice. A user experience survey was also given to the participants after the training practice.

The testing phase occurred immediately after the training session. In the testing, all three groups performed the flight tasks learned in the training activity on a Certified FTD for 1 trial (10 minutes) as a test and completed the post-training knowledge test (5 minutes) on the training activity to evaluate the training outcomes. The FTD was located in a different room than the training room but in the same building.

Results

The flight data were recorded every 0.2 seconds as a measure of training performance. After collection, data was sorted and analyzed using MS Excel and IBM SPSS. Because Chandelle is a two-phase maneuver, flight data were divided into Phase 1 data and Phase 2 data. Date sorting and coding were done after all data had been collected.

The dependent variables used for statistical analyses were flight data, including Phase 1 Bank Mean (P1Bank), Phase 1 Bank Standard Deviation (P1BankSD), Phase 1Pitch Rate Mean (P1Pitchrate), Phase 1 Pitch Rate Standard Deviation (P1PitchrateSD), Phase 2 Pitch Mean(P2 Pitch), Phase 2 Pitch Standard Deviation (P2PitchSD), Phase 2 Roll Rate (P2RollRate), Phase 2 Roll Rate Standard Deviation (P2RollRateSD), Phase 1 Turn Rate Mean (P1TurnRate), Phase 1 Turn Rate Standard Deviation (P1TurnRateSD), Phase 2 Turn Rate Mean(P2TurnRate), and Phase 2 Turn Rate Standard Deviation (P2TurnRateSD); User experience measures including, 6 UEQ subscales (Attractiveness, Perspicuity, Efficiency , Dependability, Stimulation, Novelty) , total satisfaction after training (Totalexp, Likert scale 1-5); motivation variables, including, pre-training willingness to use VR for training (WillingVR, Likert scale 1-5), pre-training willingness to use Desktop simulation for training (WillingDesk, Likert scale 1-5), post-training willingness to continue using the method used in the experiment for future training and practice (Future, Likert scale 1-5). Self-efficacy measures included, pre-training self-efficacy for the maneuver (Procedure1), post-training self-efficacy for the maneuver (Procedure2), pre-training self-efficacy for the goal of the maneuver (Goal1), post-training self-efficacy for the goal of the maneuver (Goal2), pre-training self-efficacy to perform the maneuver (Ability1), post-training self-efficacy to perform the maneuver (Ability2), and knowledge test score (Testscore). Descriptive statistics for all the tested variables per group were shown is Table 2.

Table 2

Descriptive Information for all Study Variables

Variable	Group	*N*	Mean	*SD*
P1Bank	Desktop	16	28.09	2.71
	Control	14	25.63	4.96
	VR	18	26.50	3.52
	Total	48	26.78	3.83
P1BankSD	Desktop	16	1.96	1.08
	Control	14	2.31	1.83
	VR	18	2.44	1.34
	Total	48	2.24	1.41
P1PitchRate	Desktop	16	.13	.06
	Control	14	.09	.06
	VR	18	.12	.04
	Total	48	.12	.05
P1PitchRateSD	Desktop	16	.18	.09
	Control	14	.19	.08
	VR	18	.16	.05
	Total	48	.17	.07
P2Pitch	Desktop	16	15.71	3.03
	Control	14	13.24	3.58
	VR	18	15.21	2.87
	Total	48	14.80	3.24
P2PitchSD	Desktop	16	1.84	1.03
	Control	14	2.59	1.67
	VR	18	1.61	1.09
	Total	48	1.98	1.30
P2RollRate	Desktop	16	.25	.13
	Control	14	.19	.15
	VR	18	.16	.07
	Total	48	.20	.12
P2rollRateSD	Desktop	16	.49	.22
	Control	14	.63	.16
	VR	18	.42	.16
	Total	48	.50	.20
P1TurnRate	Desktop	16	-3.62	.47
	Control	14	-3.55	.46
	VR	18	-3.45	.37
	Total	48	-3.54	.43

Variable	Group	*N*	Mean	*SD*
P1TurnRateSD	Desktop	16	1.14	.33
	Control	14	1.09	.31
	VR	18	1.03	.19
	Total	48	1.08	.28
P2TurnRate	Desktop	16	-2.25	.56
	Control	14	-2.87	.86
	VR	18	-2.19	.59
	Total	48	-2.40	.73
P2TurnRateSD	Desktop	16	1.41	.44
	Control	14	1.25	.34
	VR	18	.938	.25
	Total	48	1.18	.39
Testscore	Desktop	16	6.37	1.36
	Control	14	6.79	.80
	VR	18	7.00	1.08
	Total	48	6.73	1.12
Attractiveness	Desktop	16	1.03	1.17
	Control	14	1.86	.98
	VR	18	1.77	.88
	Total	48	1.55	1.06
Perspicuity	Desktop	16	1.29	.79
	Control	14	1.57	1.15
	VR	18	1.56	1.02
	Total	48	1.47	.98
Efficiency	Desktop	16	1.53	1.02
	Control	14	1.52	1.21
	VR	18	1.51	1.01
	Total	48	1.52	1.05
Dependability	Desktop	16	1.06	.92
	Control	14	1.70	.86
	VR	18	1.65	1.00
	Total	48	1.47	.96
Stimulation	Desktop	16	.92	1.0
	Control	14	1.59	.96
	VR	18	1.94	.72
	Total	48	1.50	.99
Novelty	Desktop	16	-.23	1.71
	Control	14	.18	1.67
	VR	18	1.61	.64
	Total	48	.58	1.59

Variable	Group	*N*	Mean	*SD*
Goal1	Desktop	16	6.44	.81
	Control	14	6.11	.62
	VR	18	6.50	.86
	Total	48	6.36	.78
Procedures1	Desktop	16	5.81	.91
	Control	14	5.10	1.24
	VR	18	5.78	1.1
	Total	48	5.59	1.13
Ability1	Desktop	16	4.44	1.55
	Control	14	4.00	1.62
	VR	18	4.61	1.38
	Total	48	4.38	1.49
Goal2	Desktop	16	6.69	.60
	Control	14	6.2	.80
	VR	18	6.67	.77
	Total	48	6.54	.74
Procedure2	Desktop	16	6.38	.50
	Control	14	5.61	1.18
	VR	18	6.28	.83
	Total	48	6.11	.91
Ability2	Desktop	16	5.28	.99
	Control	14	4.29	1.68
	VR	18	4.89	1.08
	Total	48	4.84	1.29
Totalexp	Desktop	16	3.69	.79
	Control	14	3.36	1.01
	VR	18	4.11	.76
	Total	48	3.75	.89
Future	Desktop	16	3.44	1.0
	Control	14	3.86	.66
	VR	18	4.56	.51
	Total	48	3.98	.91
Desktopwilling	Desktop	16	4.06	.93
	Control	14	4.29	.83
	VR	18	4.50	.86
	Total	48	4.29	.87
VRwilling	Desktop	16	4.50	.97
	Control	14	4.86	.36
	VR	18	4.44	1.15
	Total	48	4.58	.92

Hypothesis 1

To test H_{01}, all three groups (VR, Desktop, & Control) have no differences among the maneuver performance on the FTD , the training technique (VR, PCATD, and control group) was set as the independent variable, recorded flight data were set as the dependent variables, with the alpha-level set at .05, MANOVA tests were run among 3 Groups. Results showed significant differences in those variables, F (24, 68) = 2.08, p=.01, Wilk's Λ = .33, partial eta-squared = .42, observed power = .98. Results of the Tests of Between-Subjects Effects found significant differences in P2RollRateSD (p=0.01), P2TurnRate (p=0.01), and P2TurnRateSD (p<0.01). Post-hoc tests were then conducted to examine specific between group differences on P2RollRateSD, P2TurnRate, and P2TurnRateSD in three groups, and results were shown in Table 3.

Table 3

MANOVA Post Hoc Test Results for Hypothesis 1

Variables	Group		Mean Difference	Std. Error	Sig.
P2RollRateSD	1	2	-.14	.07	.15
		3	.07	.06	.56
	2	3	.21	.07	.01
P2TurnRate	1	2	.63	.25	.05
		3	-.06	.23	.97
	2	3	-.69	.24	.02
P2TurnRateSD	1	2	.16	.18	.47
		3	.47	.12	.01
	2	3	.32	.12	.05

Note: Group 1= Desktop, 2= Control, 3= VR

Post-hoc testing suggested that the VR group P2RollRateSD (M= .425, SD= .16) was significantly lower than the Control group P2RollRateSD (M= .63, SD= .168), the VR group

P2TurnRate (*M*=2.19, *SD*=.59) and Desktop group P2TurnRate(*M*=2.25, *SD*=.56) were both significantly lower than the Control Group P2TurnRate(*M*=2.88, *SD*=.86), and the VR group P2TurnRateSD(*M*=.94, *SD*=.25) was significantly lower than both the Control Group P2TurnRateSD(*M*=1.25 *SD*=.34) and the Desktop P2TurnRateSD(*M*=1.41, *SD*=.44). The lower value of those variables indicates a smoother performance. This pattern of results suggested that VR group out performed the Desktop and Control group, therefore, H_{01} was rejected.

Hypothesis 2

To test H_{02}: All three groups have no difference among the scores on knowledge tests , the training technique (VR, Desktop, and control group) was set as the independent variable, and the post-training Chandelle maneuver knowledge test score was set as the dependent variable. With the alpha-level set at .05, a one-way between-subjects ANOVA test was not significant, $F(2, 45) = 1.352$, $p = 0.269$. The results suggested retaining the H_{02}.

Hypothesis 3

To test H_{03}: User experience among all three groups has no difference, the training technique (VR and PCATD group) was set as the independent variable, and the User experience Survey results were set as the dependent variables. The Independent-sample T-tests were run between the VR and Desktop group with alpha-level set at .05. Test results suggested that VR Group Attractivene (*M*=1.77, *SD*= .88) was significantly higher than Desktop Group (*M*=1.03, *SD*=1.17), $t(32)=-2.09$, p=.04. Attractivene is a subscale of UEQ, which represents the overall impression of the product to the users. VR Group Stimulation (*M*=1.94, *SD*=.01) was significantly higher than Desktop Group (*M*=.92, *SD*= 1.07). $t(32)= -3.31$, p= .04. Stimulation is a subscale of UEQ, which represents how exciting and motivating it is to use the interaction. VR Group Novelty (*M*=1.61, *SD*=.64) was significantly higher than Desktop Group(*M*=-2.34, *SD*=1.71), *t*(18.69) = -4.08, *p*< .01. Novelty represents how innovative and creative the interaction is. Therefore, the H_{03} was rejected.

Hypothesis 4

To test H_{04}: The trainees' motivation to use VR flight simulation for training and learning purposes is not different than motivation for use of conventional Desktop Flight Simulation training techniques or training techniques used in the control group. Paired T-tests were run on the responses to the questions about participants' willingness to use VR and Desktop flight simulation for learning. An ANOVA test was run on the response to questions about users' willingness to use the training method for future training in the User experience Survey. Paired T-tests showed that participants' willingness to use VR for training and learning (*M*=4.85, *SD*=.92) was significantly higher than to use Desktop simulation (*M*=4.29, *SD*=.87), *t*(47)=-2.19, *p*= .033. ANOVA test results suggested that user's willingness to use VR techniques (*M*=4.56, *SD*=.51) for future training was significantly higher than using the Desktop simulation (*M*=3.44, *SD*=1.09) or the Control group method(*M*=3.86, *SD*= .66), *F*(2, 45)= 8.72, *p*<0.01. Therefore, H_{04} was rejected. Post Hoc test results showed in Table 4.

Table 4

ANOVA test Post Hoc Test Results for Hypothesis 4

Varibales	Group		Mean Difference	Std. Error	Sig.
Willingness to use current training method for future learning and practice	Desktopp	Control	-.42	.33	.42
		VR	-1.12	.29	.01
	Control	VR	-.69	.21	.01

Hypothesis 5

To test $H_{05:}$ User experience is not related to the maneuver performance on the FTD. User experience Survey results and recorded flight parameter data were set as variables for correlation coefficient tests. Results found that P2RollRate was positively correlated with Novelty, $r(48) = .29$, $p = .04$. P1TurnRateSD was negatively correlated with Attractiveness, $r(48)=-.424$, $p< .01$, Perspicuity, $r(48)= -.40$, $p< .01$, Efficiency, r(48)= -.32, $p= .03$, Stimulation, $r(48) =-.33$, $p=.02$, and Novelty, $r(48)= -.30$, $p=.04$. P2 TurnRate was positively correlated with Attractiveness, $r(48) =.31$, $p= .04$, Efficiency, $r(48)= .35$, $p= .02$, and Dependability, $r(48) =.43$, $p <.01$. Because the lower value of standard deviation variables indicates a smoother performance, the negative correlation actually showed that user experience was positively related to the maneuver performance. Therefore, H_{05} was rejected.

Hypothesis 6 and 7

To test $H_{06:}$ the participants' self-efficacy pertaining to the trained flight maneuver has no changes after practice in all three groups, and H_{07}: there is no difference in post-training self-efficacy on the selected flight maneuver among the three groups. Time (pre-practice & post-practice,) and the training technique (VR, Desktop, and Control) were set as the independent variable, and Self-efficacy was set as the dependent variable. Repeated-measures MANOVA tests were run on responses to the self-efficacy questionnaire. Results of the MANOVA found significant within-group differences in self-efficacy scores, for self-efficacy pertaining to the goal of the maneuver, $F(1, 45) = 5.16$, $p=.03$, Wilk's $\Lambda = .90$, partial eta-squared = .10, observed power = .61, for self-efficacy pertaining to the procedure of maneuver, $F\ (1, 45)=13.20$, $p<.01$, Wilk's $\Lambda = .77$, partial eta-squared = .23, observed power = .95, and for self-efficacy pertaining to perform the maneuver, $F(1, 45)=7.67$, p=.01 Wilk's $\Lambda = .85$, partial eta-squared = .15, observed power = .77. No significant differences were found between groups or in the interactions. Therefore, H_{06} was rejected and H_{07} was retained.

Hypothesis 8

To test $H_{08:}$ The participants' self-efficacy pertaining to the trained flight maneuver is not related to the maneuver performance. Pre-practice and post-practice self-efficacy questionnaire responses, and recorded flight parameter data were set as variables for correlation coefficient tests. Results found that pre-practice self-efficacy for the procedure of the maneuver was significantly correlated to P2Pitch, $r(48)=.48$, $p=.01$, and P2RollRate, $r(48)=.31$, $p=.034$. Post-practice self-efficacy to understanding the procedure of maneuver was significantly correlated to P1Bank, r(48)= .38, $p=.01$, P1Pitch, $r(48)= .38$, $p=.01$, P2Pitch, $r(48)=.41$, $p< .01$, and P1TurnRate, $r(48)=.30$, p=.04. Pre-practice self-efficacy pertaining to performing the maneuver was significantly correlated with P2Pitch, $r(48) =.38$, $p= .01$, and P2TurnRate, $r(48) =.40$, $p<.01$. Post-practice self-efficacy pertaining to performing the maneuver was significantly correlated P1Pitch, $r(48)=.31$, $p= .03$, P2Pitch, $r(48)= .30$, $p= .04$, and P2TurnRate, $r(48)= .33$, $p=.02$. Therefore, H_{08} was rejected.

Discussion

The current project focused on determining the effectiveness of VR flight simulation training and if it would be more effective than traditional PCATD- based training. Furthermore, the researcher investigated user experience for the VR flight simulation, as well as whether such technology would improve the affective outcomes of training that enhance student pilots' learning ability and motivation to gain new knowledge and skills.

VR for Procedural Learning

Hypothesis 1 was focused on testing the difference between procedural learning outcomes of different training techniques. Significant differences in maneuver performance on the FTD among three groups were detected, specifically, the VR group performed significantly better than the other groups in the second phase of the Chandelle maneuver (second 90° of the turn) on the FTD. Therefore, H_{01} was rejected. During first phase of Chandelle (first 90°), the polit needs to set at a fixed angle of bank at 30°, increase power, and increase pitch attitude at a rate such that maximum pitch-up occurs at the completion of the first 90°. In the second phase of the maneuver, the pilots needed to continue rollout the plane, and the vertical component of lift increase. As the airspeed continues to decrease, a slight increase of elevator back pressure is required to keep the pitch attitude from decreasing which requires the pilot utilizing a higher level of coordination of the flight controls comparing to first phase of the maneuver. The current study results echoed supported prior research findings. Bowman et al (2009) suggested that a higher level of visual fidelity provides better performance in procedure memorization tasks, and Kwon's study (2019) found that high fidelity VR simulation helps users to recognize a virtual experience as an experience that could help them to gain a better understanding of the subject learned during the simulation. Compared to conventional desktop flight simulation, VR flight simulation provides higher visual fidelity with 360 degree of dynamic field of view (FOV) instead of conventional desktop simulation’s 180 degree, fixed field of view. For complex flight training content like the Chandelle, which generally requires pilots to look around, define out of aircraft visual reference point, and to adjust flight control coordinating to the aircraft position, VR simulation's 360 degree of dynamic field of view definitely could provide better visual cues than conventional desktop flight simulation. The advantages of high visual fidelity in VR simulation might give the participants a better sense of spatial orientation and provide a more realistic training experience, which helped the VR group to gain a better procedural skill than the other two groups.

VR Simulation and Cognitive Learning

Several studies have been done to explore the effect of VR technology on learning; some of the those studies showed that VR facilitated cognitive learning, while others did not. Makransky et al. (2019) found that VR was more effective in addressing problems in a real-world lab-setting than a text lesson. Kozhevnikov et al. (2013) found that students who viewed a lesson on a VR display performed significantly better on transfer tests than students who learned using a desktop display. In contrast, Menin et al. (2022), found that VR promoted procedural but not conceptual learning in fire safety training. Parong et al. (2021) reported immersive VR created high emotional arousal, however, it also caused a cognitive distraction, which lead to less learning. One interesting common finding of all the VR related studies, including the current study, is that learners’ motivation to use VR technology for studying is higher than using other mediums.

Hypothesis 2 of the current study aimed to see if VR technology has a faciliatory effect on cognitive learning. A comparison of the three groups’ post-training knowledge test scores did not show significant differences among groups. Therefore, Hypothesis 2 was rejected. A limitation in the present study was related to participants' unfamiliarity with using VR simulation for learning purposes. While students were able to use VR for procedural skills and saw the alignment between VR and actual flight training, they have not been taught to think about or apply procedural VR skill acquisition to cognitive and non-procedural knowledge acquisition. Perhaps, different results may have been found if we extended the experiment period and increased the sessions of training, thereby increasing opportunities for that link to develop. For the current study, most participants had little experience with VR flight simulation; eighty percent of participants had zero VR flight simulation experience, besides using the VR simulation to learn flight maneuvers, and there is also a learning curve for the users to become proficient in using the VR flight simulation itself. A focus on skill development then limited any higher order cognitive application of the training. The true relationship between VR technology and its effect on cognitive learning, then, is still unknown. Further research could incorporate a longer experimental period and consider participants' expertise in VR systems as variables to explore in an examination of the effect VR training may have on cognitive learning outcomes.

VR Flight Simulation User Experience and Motivation

Hypotheses 3 and 4 focused on user experience and motivation for VR flight simulation. The UEQ were used as to measure user experience and customized questions were used as to measure user motivation. The UEQ scores of VR and Desktop groups were shown in Table 5.

Table 5

UEQ scores

	Group	*N*	*M*	*SD*
Attractiveness	Desktop	16	1.03	1.17
	VR	18	1.77	.88
Perspicuity	Desktop	16	1.30	.79
	VR	18	1.56	1.02
Efficiency	Desktop	16	1.5	1.02
	VR	18	1.51	1.01
Dependability	Desktop	16	1.06	.92
	VR	18	1.65	1.00
Stimulation	Desktop	16	.92	1.07
	VR	18	1.94	.72
Novelty	Desktop	16	-.23	1.71
	VR	18	1.6	.64

In a comparison of those in the VR flight simulation group score to UEQ benchmark scores (Schrepp, 2017), which are based on data from user tests of 452 products, the VR flight simulation group's mean Attractiveness score was good (above 75% results), Perspicuity was above average (above 50% of results), Efficiency was good (above 75% results), Dependability was good (above 75% results), Stimulation was excellent (in the range of the 10% best results), and Novelty was also excellent (in the range of the 10% best results). T-tests between Desktop Simulation and VR simulation showed that the VR flight simulation group had significantly higher scores than the Desktop simulation in Attractiveness, Stimulation, and Novelty. Hypothesis 3 and 4 were both retained.

Attractiveness of UEQ represents the overall impression of the product. The VR flight simulation's Attractiveness was good (above 75% results) compared to the UEQ benchmark and higher than desktop simulation. This means participants generally liked to use the VR flight simulation and found it more attractive than using the conventional desktop simulation.

Perspicuity of UEQ represents how easy it is to learn to use the product. The VR flight simulation group's Perspicuity was above average (above 50% of results) compared to the UEQ benchmark and not significantly different from the desktop simulation. This means participants need to spend the same time and effort to learn how to use the VR simulation as desktop simulation, and the difficulty to learn VR is at a similar level to learning a desktop flight simulation.

Efficiency of UEQ represents the likelihood that users can solve their tasks without unnecessary effort. The VR flight simulation's Efficiency was good (above 75% results) compared to the UEQ benchmark and not significantly different from the desktop simulation. This means that participants can perform their tasks fast in a pragmatic way with the VR simulation and could be as efficient as using the Desktop simulation.

Dependability of UEQ represents the degree of the user's control over the product. The VR flight simulation's Dependability was good (above 75% results) compared to the UEQ benchmark and not significantly different from the desktop simulation. This means the interaction that VR flight simulation provided was predictable and meets the participants' expectation for flight training in the same manner as a desktop simulation.

Stimulation of UEQ represents how interesting, exciting, and motivating it is to use the product. The VR flight simulation group's Stimulation score was excellent (in the range of the 10% best results) compared to the UEQ benchmark and significantly higher than the desktop simulation group's mean score. A similar trend was found in the responses to questions about participants' willingness to use VR and desktop flight simulation for learning and the response to questions about users' willingness to use the experimental training method for future training. Also, the comments provided by the participants after the experiment showed that most pilots have an interest in using VR flight simulation for flight training, regardless of whether they actually received the training in VR. Those findings were consistent with prior research (Makransky, et al. 2019; Kwon, 2019 ; Menin, et al. 2022). We can see that participants generally feel using VR flight simulation is more interesting and motivating than using a desktop simulation. Perhaps, the participant may find the features of the VR flight simulation, such as the realistic 3D environment, 360 degree of dynamic field of view, and the detailed interactive interface could be beneficial to training, or that the VR interaction is more close to the real flight experience. Another factor that contributed to the results could be that participants have extra excitement to try and use new technology.

Novelty of UEQ represents how innovative and creative the product is. The VR flight simulation group's Novelty score was excellent (in the range of the 10% best results) compared to the UEQ benchmark and significantly higher than the desktop simulation group score. Today, VR technology is still relatively new to most populations, so it's understandable that most student pilot participants feel the VR flight simulation is more innovative and creative than a desktop flight simulation, which has been around for decades.

Depending on the current study results, we can see a pattern that participants generally like to use VR flight simulation, and their overall user experience is in some ways higher than using a desktop simulation. With potentially better user experience, VR flight simulation could be a valuable alternative to conventional desktop flight simulation as a training and learning tool.

User Experience and Training Results

Hypothesis 5 centered on the correlation between the user experience and the maneuver performance. Hypothesis testing results found that each of the 6 dimensions of UEQ were significantly related to one or multiple of the flight parameter variables, and Hypothesis 5 was reject. In addition to Hypothesis 5, correlation coefficient tests were also run between UEQ scores and the post-training knowledge test. Significant positive correlations were found between UEQ scores and knowledge with Dependability, $r(48)=.34$, $p= .02$, Stimulati on, $r(48)=.34$, $p=.02$, and Novelty, $r(48)= .30$, $p= .04$ significantly correlated with knowledge. These findings support the idea that products with a higher user experience could be more beneficial for acquiring knowledge and skills. Better user experience could also generate more interest and motivation to continually use the product for learning. Based on the study results, there was sufficient evidence to support a theory that the user experience of flight simulation is positively related to its training results. It’s predictable that, with the advancement of technology and design, VR flight simulation could provide a better user experience for the pilots and be a more efficient training tool. This is a potentially valuable area for future research.

Self-efficacy and Training Results

Hypotheses 6 to 8 centered on the effect of different training methods on self-efficacy, and the correlation between self-efficacy and training results. Hypothesis 6 proposed that the participants’ self-efficacy pertaining to the trained flight maneuver has no changes after practice in all three groups; Hypothesis 7 proposed that there is no difference in post-training self-efficacy on the selected flight maneuver among the three groups. Hypotheses 6 and 7 were assessed using Repeated-measures MANOVA tests. Results of the MANOVA test found significant within-group differences between pre-training and post-training self-efficacy scores, however, no between-group differences were found. Self-efficacy is an individual's belief in her or his capability to complete a task, and it affects how people approach challenges and reach goals (Bandura, 1994). Based on the participants' responses, all three groups showed increases in self-efficacy for their understanding of the goal and procedure of the Chandelle, as well as the ability to perform the maneuver; all participants felt they gained skills and were more confident about the training after their training session. There was no evidence of a difference in post-practice efficacy levels among the three groups. These findings are consistent with Buttussi & Chittaro's (2018) and Reweti et al.'s (2017) findings on the effects of different types of training techniques on self-efficacy. Bandura (1994) suggested that a major way to increase self-efficacy is to gain mastery experiences in performing the given behavior. Understandably, all three groups had similar post-practice efficacy levels, because they all had experience in practicing the same flight maneuver.

Hypothesis 8 proposed that the participants' self-efficacy on the trained flight maneuver was not related to the maneuver performance on FTD. Pre-practice and post-practice self-efficacy questionnaire responses, and recorded flight parameter data were set as variables for correlation coefficient tests. Hypothesis testing results showed that both pre-practice and post-practice self-efficacy levels were significantly, and positively correlated to one or multiple of the flight parameter variables. Hypothesis 8 was rejected. The current study findings were again consistent with Social Cognitive Theory, that higher self-efficacy has a positive effect on performance (Bandura, 1999).

Current findings supported that VR flight simulation has a positive effect on self-efficacy pertaining to training activities, and VR is as effective as desktop flight simulation for training purposes in terms of self-efficacy gain.

Limitations and Future Work

This study had multiple limitations resulting from the participants, the technology, and the study design. The first limitation was the number of participants used in this study. The sample size of the current study was small and only included 48 participants, which just reached the pre-study power analysis that determined the minimum required number to obtain an effect size of 0.3 with a confidence interval of 0.95. Most participants were student pilots with PPL or in the middle of PPL training (passed SOLO) from the same Aeronautical University. Different results might be found if using other level pilots or recruiting from other regions. In addition, 83.3% of the participants were male, so the sample could be gender-biased and the results less applicable to female flight students. The findings could also be susceptible to selection bias, as individuals who volunteered to participate may have had an extra interest in flight simulation training, and the extra interest may have generated more positive results to confirm their personal bias. Last, participants may have felt undue pressure to provide positive feedback on the training session. These constraints may make the findings being less generalizable than studies involving a larger more randomized sample.

The second global limitation was the technology used in this study. Due to resource constraints, the experiments were conducted via HTC VIVE pro VR set with X-plane 11 flight simulation software. The current VR flight simulation configuration used in this study may not be the best solution to provide the most effective training experience to the participant. Another limitation related to the participants' unfamiliarity with either the VR flight simulation or the desktop flight simulations used in this study. The duration of the training activity was relatively short, only 45 minutes to 60 minutes, and the duration may not have been enough for the participants to fully learn how to use the VR flight simulation or desktop flight simulation effectively. The novelty of the technology and unfamiliarity with the control devices could cause negative consequences for learning. Future research may consider having sufficient resources to familiarize participants with the flight simulation technologies before a training session.

The third limitation was the training content and study design. This project only focuses on a single, short training session on the Chandelle maneuver. The effectiveness of the experimental results could only be used to explain the results of this procedure. Different training contexts could influence the effect of the training techniques on training results. Pilot training is a long-term process, therefore, it's important to explore the effect of VR simulation on training results on a long-term scale.

To determine the true effects of VR flight simulation on pilot training results, future work is needed to examine the validity of the current finding with a long-term training intervention with a larger randomized sample. Future studies can also examine the effect of newer types of VR flight simulation using more comprehensive training contexts, as well as explore the relationship between participants' expertise in VR technology and training results

Conclusion

The present study compared the effectiveness of VR-based flight training to traditional PCATD- based training methods. There has been little research done in this area, and the present study was a pioneer study in the field. Based on a quasi-transfer of training study design, the researchers found that participants performed better in the post-training maneuver performance in the VR group. Findings also supported that VR flight simulation could provide a better user experience and generate a higher motivation for usage.

Despite the limitations, the present study findings contributed positive evidence that VR flight simulation has a large potential to be an effective flight training tool. Predictably, the advancements in VR technology will provide opportunities for research examining how VR flight training applications affect pilot skill development in the future.

References

Allen, J.A., 1986, Maintenance training simulator fidelity and individual difference in transfer of training, *Human Factors*, *28*(5), 497–509.

Anton, D., Kurillo, G., & Bajcsy, R. (2018). User experience and interaction performance in 2D/3D telecollaboration. *Future Generation Computer Systems*, *82*, 77–88. https://doi.org/10.1016/j.future.2017.12.055

Atkins, R. J., Landsdowne, A. T. G., Pfister, H. P., & Provost, S. C. (2002). Conversion between control mechanisms in simulated flight: An ab initio quasi-transfer study. *Australian Journal of Psychology*, 54(3), 144–149. https://doi.org/10.1080/00049530412331312714

Baarspul, M. (1990). A review of flight simulation techniques. *Progress in Aerospace Sciences*, 27(1), 1–120. https://doi.org/10.1016/0376-0421(90)90006-6

Bandura, A. (1994). Self-efficacy. *Encyclopedia of Human Behavior*, 4, 71-81.

Bandura, A. (1999). Social cognitive theory: An agentic perspective. *Asian Journal of Social Psychology*, 2, 21–41.

Beckman, W. S. (2000). A Comparison of the Effectiveness of PC-Based Aviation Training Devices and Conventional Flight Training Devices for Instrument Flight Training. *Journal of Aviation/Aerospace Education & Research*, 9(2).

Beckman, W. (2003). A Evaluation of the Perceived Effect of Personal Computer Aviation Training Devices for Instrument Flight Training. *Journal of Aviation/Aerospace Education & Research*, 12(3). https://doi.org/10.15394/jaaer.2003.1568

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7), 36–43. https://doi.org/10.1109/MC.2007.257

Bowman, D. A., Sowndararajan, A., Ragan, E. D., & Kopper, R. (2009). Higher levels of immersion improve procedure memorization performance. Proceedings of the Joint Virtual Reality Conference of EGVE - The 15th Eurographics Symposium on Virtual Environments, ICAT, EuroVR 2009, 121–128. https://doi.org/10.2312/EGVE/JVRC09/121-128

Brennesholtz, M. S. (2018). 3-1: Invited Paper : VR Standards and Guidelines . SID Symposium Digest of Technical Papers, 49(1), 1–4. https://doi.org/10.1002/sdtp.12476

Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology* (2nd ed.). Hoboken, NJ: Wiley-Interscience.

Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1063–1076. https://doi.org/10.1109/TVCG.2017.2653117

Caro, P.W., 1988, Flight training and simulation, in Human Factors in Aviation, Wiener, E.L. and Nagel, D.C., Eds., San Diego, CA: Academic Press.

Chang, C. W., Yeh, S. C., Li, M., & Yao, E. (2019). The Introduction of a Novel Virtual Reality Training System for Gynecology Learning and Its User Experience Research. IEEE Access, 7, 43637–43653. https://doi.org/10.1109/ACCESS.2019.2905143

Chen, G., Lu, X., Liao, D., & Chen, Q. (2008). Tactical communication simulation training system based on VR. Proceedings - 2008 International Conference on Advanced Computer Theory and Engineering, ICACTE 2008, 494–498. https://doi.org/10.1109/ICACTE.2008.30

Combs, A. (2001). The PCATD's Role in the Cognitive Processes of Flight Training. *Journal of Aviation/Aerospace Education & Research*, *10*(3).

Dahlström, N. (2008). Pilot training in our time - use of flight training devices and simulators. *Aviation*, *12*(1), 22-27. doi:10.3846/1648-7788.2008.12.22-27

Ertmer, P.A. and Newby, T.J. (2013), Behaviorism, Cognitivism, Constructivism: Comparing Critical Features From an Instructional Design Perspective. *Perf. Improvement Qrtly*, *26*: 43-71. doi:10.1002/piq.21143

Federal Aviation Administration. (2016). Airplane Flying Handbook (FAA-H-8083-3B). Washington D.C.: U.S. Department of Transportation.

Federal Aviation Administration. (2018). FAA approval of aviation training devices and their use for training and experience. (Advisory Circular 61-136B). Washington D.C.: U.S. Department of Transportation.

Gavan Lintern , Henry L. Taylor , Jefferson M. Koonce , Robert H. Kaiser & Gregory A. Morrison (1997) Transfer and Quasi-Transfer Effects of Scene Detail and Visual Augmentation in Landing Training, *The International Journal of Aviation Psychology*, 7:2, 149-169.

Geršak, G., Lu, H., & Guna, J. (2018). Effect of VR technology matureness on VR sickness. *Multimedia Tools and Applications*, (November). https://doi.org/10.1007/s11042-018-6969-2

Gheorghiu, A. (2013). Flight Simulation in Air Force Training. a Knowledge Transfer Eficiency Perspective. *Journal of Defense Resources Management* (JoDRM), 4(2), 153–158.

Grabowski, A.,& Jankowski, J. (2015). Virtual Reality-based pilot training for underground coal miners. *Safety Science*, 72, 310-314.

Greunke, L., & Sadagic, A. (2016). Taking Immersive VR Leap in Training of Landing Signal Officers. *IEEE Transactions on Visualization and Computer Graphics*, 22(4), 1482–1491. https://doi.org/10.1109/TVCG.2016.2518098

Homan, W. J., & Williams, K. D. (1997). Effects of PC-Based Pretraining on Pilots' Performance in an Approved Flight-Training Device. *Journal of Aviation/Aerospace Education & Research*, 7(3).

Jones, E.R., Hennessy, R.T., and Deutsch, S., 1985, Human Factors Aspects of Simulation, Washington D.C.National Academy Press.

Jorna, P.G.A.M. (1993). Heart rate and workload variations in actual and simulated flight. *Ergonomics*, 36(9), 1043-1054.

Joyce, R. D., & Robinson, S. K. (2015). The Rapidly Reconfigurable Research Cockpit. (June), 1–11. https://doi.org/10.2514/6.2015-2336

Kennedy, R. S., Lane, N. E., Kevin, S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire : An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3), 203–220. https://doi.org/10.1207/s15327108ijap0303

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The fourlevels* (Third ed.). San Francisco, CA: Berrett-Koehler

Koonce, J. M., & Bramble Jr, M. J. (1998). Personal Computer- Based Flight Training Devices. *The International Journal of Aviation Psychology*, 8(3), 223–242. https://doi.org/10.1207/s15327108ijap0803

Kozhevnikov, M., Gurlitt, J., & Kozhevnikov, M. (2013). Learning Relative Motion Concepts in Immersive and Non-immersive Virtual Environments. *Journal of Science Education and Technology*, 22(6), 952–962. https://doi.org/10.1007/s10956-013-9441-0

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of Cognitive, Skill-Based, and Affective Theories of Learning Outcomes to New Methods of Training Evaluation. *Journal of Applied Psychology*, *78*(2), 311–328. https://doi.org/10.1037/0021-9010.78.2.311

Kwon, C. (2019). Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virtual Reality*, *23*(1), 101–118. https://doi.org/10.1007/s10055-018-0364-1

LaViola Jr, J. (2000). A Discussion of Cybersickness in Virtual Environments. ACM SIGCHI Bulletin, 32(1), 23–26.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. USAB 2008: *HCI and Usability for Education and Work*, pp 63-76 https://doi.org/10.1007/978-3-540-89350-9_6

Lee, A.T., (2009). Flight simulation virtual environments in aviation. Burlington,22

Lintern, G., Taylor, H. L., Koonce, J. M., Kaiser, R. H., & Morrison, G. A. (1997). Transfer and quasi-transfer effects of scene detail and visual augmentation in landing training. *International Journal of Aviation Psychology*, 7(2), 149–169. https://doi.org/10.1207/s15327108ijap0702_4

Liu, D., Blickensderfer, E., Vicenzi, D. & Macchiarella D (2008)., Transfer of training, Ch3, in Vincenzi D. and Wise J. (eds), *Human Factors and Simulation*. CRC Press, Taylor & Francis Group.

Liu, D., Vicenzi, D. & Macchiarella D (2008), Simulation fidelity, Ch 4, in Vincenzi D. and Wise J. (eds), *Human Factors and Simulation* CRC Press, Taylor & Francis Group.

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, 35(6), 691–707. https://doi.org/10.1111/jcal.12375

Maytin, M., Daily, T. P., & Carillo, R. G. (2015). Virtual reality lead extraction as a method for training new physicians: A pilot study. PACE - *Pacing and Clinical Electrophysiology*, 38(3), 319–325. https://doi.org/10.1111/pace.12546

Mccarty, W. D., Sheasby, S., Amburn, P., Stytz, M. R., & Switzer, C. (1994). A virtual cockpit for a distributed interactive simulation. *IEEE Computer Graphics and Applications*, 14 (1), 49-53.

Mcdermott, J. T. (2005). *A comparison of the effectiveness of a personal computer-based aircraft training device and a flight training device at improving pilot instrument proficiency: A case study in leading regulatory change in aviation education.* [Doctoral dissertation, Bowling Green State University]. ProQuest Dissertations and Theses.

Menin, A., Torchelsen, R., & Nedel, L. (2022). The effects of VR in training simulators: Exploring perception and knowledge gain. *Computers and Graphics* (Pergamon), 102, 402–412. https://doi.org/10.1016/j.cag.2021.09.015

Mu¨hlberger, A., Herrmann, M. J., Wiedemann, G., Ellgring, H., & Pauli, P. (2001). Repeated exposure of flight phobics to flights in virtual reality. *Behaviour Research and Therapy*, *39*(9).

Myers, P. L., Starr, A. W., & Mullins, K. (2018). Flight simulator fidelity, training transfer, and the role of instructors in optimizing learning. *International Journal of Aviation, Aeronautics, and Aerospace*, *5*(1). https://doi.org/10.15394/ijaaa.2018.1203

Noble, C. (2002). The relationship between fidelity and learning in aviation training and assessment. *Journal of Air Transportation World Wide, 11*, 33-54.

Norris, M., Spicer, K., & Byrd, T. (2019). Virtual Reality: The new pathway for effective safety training. *Professional Safety*, *64*(06), 36–39.

Oberhauser, M., & Dreyer, D. (2017). A virtual reality flight simulator for human factors engineering. *Cognition, Technology and Work*, 19(2–3), 263–277. https://doi.org/10.1007/s10111-017-0421-7

Oberhauser, M., Dreyer, D., Braunstingl, R., & Koglbauer, I. (2018). What's Real About Virtual Reality Flight Simulation?: Comparing the Fidelity of a Virtual Reality With a Conventional Flight Simulation Environment. *Aviation Psychology and Applied Human Factors*, *8*(1), 22–34. https://doi.org/10.1027/2192-0923/a000134

O'Malley, S., Rajagobal, A., Grundy, J. G., Mohrenshildt, M. V., & Shedden, J. M. (2016). Exposure to Disturbance Motion During Practice in an Analog of a Flight Task Influences Flight Control of Naive Participants. *International Journal of Aviation Psychology*, 26(3–4), 63–74. https://doi.org/10.1080/10508414.2017.1295246

Oritz, G. (1993). Transfer of Learning Effectiveness: PC-Based Flight Simulation. *Journal of Aviation/Aerospace Education & Research*, *3*(2). https://doi.org/10.15394/jaaer.1993.1097

Page, R. (2000). Brief history of flight simulation. SimTecT 2000 Proceedings, 1–11. https://doi.org/10.1.1.132.5428

Parong, J., & Mayer, R. (2018). Learning Science in Immersive Virtual Reality Article in Journal of Educational Psychology. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1037/edu0000241

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning*, *37*(1), 226–241. https://doi.org/10.1111/jcal.12482

Reweti, S., Gilbey, A., & Jeffrey, L. (2017). Efficacy of low-cost PC-based aviation training devices. *Journal of Information Technology Education: Research*, *16*(1), 127–142. https://doi.org/10.28945/3682

Robertson, A., & Zelenko, M. (2014). Voices from the past: An oral history of a technology whose time has come again. Retrieved from http://www.theverge.com/a/virtual-reality/oral_history

Rothbaum, B. O., Anderson, P., Zimand, E., Hodges, L., Lang, D., & Wilson, J. (2006). Virtual Reality Exposure Therapy and Standard (in Vivo) Exposure Therapy in the Treatment of Fear of Flying. *Behavior Therapy*, *37*(1), 80–90. https://doi.org/10.1016/j.beth.2005.04.004

Roscoe, S.N. (1980). Aviation psychology. Ames, IA: The Iowa State University Press

Sacks, R., Perlman, A., & Barak, R. (2013). Construction safety training using immersive virtual reality. *Construction Management and Economics*, *31*(9), 1005–1017. https://doi.org/10.1080/01446193.2013.828844

Sanchez-Vives, M.V. and Slater, M. (2005) From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4), 332–9.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence*, *4*(4), 40. https://doi.org/10.9781/ijimai.2017.445

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale, *Measures in health psychology: A user's portfolio*. J. Weinman, S. Wright, & M. Johnston, Eds, 35-37.

Somrak, A., Humar, I., Hossain, M. S., Alhamid, M. F., Hossain, M. A., & Guna, J. (2019). Estimating VR Sickness and user experience using different HMD technologies: An evaluation study. *Future Generation Computer Systems*, 94, 302–316. https://doi.org/10.1016/j.future.2018.11.041

Sousa, D. (2017). How the brain learns (5th ed.). Thousand Oaks, CA: Corwin.

Spector, P.E., 2003, Industrial/Organizational Psychology: Research and Practice, 3rd ed., New York: John Wiley & Sons

Su, K. W., Chen, S. C., Lin, P. H., & Hsieh, C. I. (2019). Evaluating the user interface and experience of VR in the electronic commerce environment: a hybrid approach. *Virtual Reality*. https://doi.org/10.1007/s10055-019-00394-w

Taber, M. J. (2014). Simulation fidelity and contextual interference in helicopter underwater egress training: An analysis of training and retention of egress skills. *Safety Science*, 62, 271-278. doi:10.1016/j.ssci.2013.08.019

Talleur, D. A., Taylor, H. L., Emanuel, T. W., Rantanen, E., & Gary L. Bradshaw. (2003). Personal Computer Aviation Training Devices: Their Effectiveness for Maintaining Instrument Currency. *The International Journal of Aviation Psychology*, 13(4), 387–399. https://doi.org/10.1207/S15327108IJAP1304

Taylor, H.L., Lintern ,G., Hulin, C.L., Talleur, D.A. , Emanuel Jr. T.W. & Phillips, S. I. (1999) Transfer of Training Effectiveness of a Personal Computer Aviation Training Device, *The International Journal of Aviation Psychology*, 9 (4), 319-335.

Taylor, H. L., Lintern, G., & Koonce, J. M. (1993). Quasi-transfer as a predictor of transfer from simulator to airplane. *Journal of General Psychology*, Vol. 120, pp. 257–276. https://doi.org/10.1080/00221309.1993.9711147

Thomas, R. L. (2018). *Effect of Active Learning on Instrument Rated Pilots' Knowledge and Self-efficacy*. [Doctoral dissertation, Embry-Riddle Aeronautical University]. ProQuest Dissertations and Theses.

Thompson, N.T, Carroll, M.B & Deaton, J.E. (2008). Justification for Use of Simulation, Ch 2, in Vincenzi D. and Wise J. (eds), Human Factors and Simulation CRC Press, Taylor & Francis Group.

University of North Dakota (2021, March 2), *Commercial Chandelles* [Video]. Youtube. https://www.youtube.com/watch?v=Ml8YI7oj2Q8&t=533s

Vaden, E.A & Hall,S. (2005). The Effect of Simulator Platform Motion on Pilot Training Transfer: A Meta-Analysis, *The International Journal of Aviation Psychology*, *15*(4), 375-393.

Vankipuram, M., Kahol, K., McLaren, A., & Panchanathan, S. (2010). A virtual reality simulator for orthopedic basic skills: A design and validation study. *Journal of Biomedical Informatics*, 43(5), 661–668. https://doi.org/10.1016/j.jbi.2010.05.016

Verdaasdonk, E. G. G., Dankelman, J., Lange, J. F., & Stassen, L. P. S. (2008). Transfer validity of laparoscopic knot-tying training on a VR simulator to a realistic environment: A randomized controlled trial. *Surgical Endoscopy and Other Interventional Techniques*, 22(7), 1636–1642. https://doi.org/10.1007/s00464-007-9672-3

Yavrucuk, I., Kubali, E., & Tarimci, O. (2011). A low cost flight simulator using virtual reality tools. *IEEE Aerospace and Electronic Systems Magazine*, 26(4), 10–14. https://doi.org/10.1109/MAES.2011.5763338

Yu, M., Zhou, R., Wang, H., & Zhao, W. (2019). An evaluation for VR glasses system user experience: The influence factors of interactive operation and motion sickness. *Applied Ergonomics*, 74(August 2018), 206–213. https://doi.org/10.1016/j.apergo.2018.08.012

Williges, B. H., Roscoe, S. N. & Williges, R. C., (2001). Synthetic flight training revisited transfer of training and the measurement of training effectiveness. In R. Sweezey, & D. Andrews (Eds.), Readings in training and simulation: a 30-year perspective (pp. 2-12). Santa Monica, CA:Human Factors and Ergonomics Society

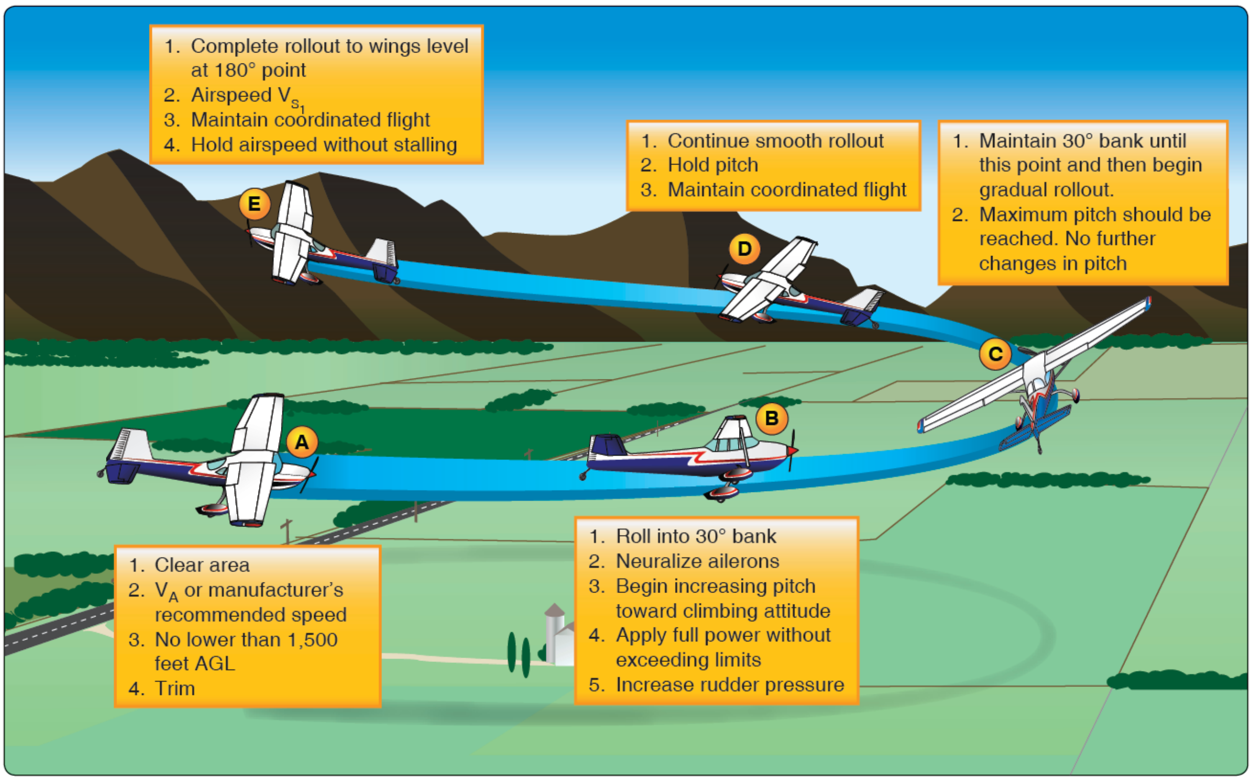
Zaa, P. M. T., Schroeder, J. A., & Chung, W. W. (2015). Transfer of training on the vertical motion simulator. *Journal of Aircraft*, 52(6), 1971–1984 https://doi.org/10.2514/1.C033115

Appendix A

Chandelle (Adapted from FAA Airplane Flying Handbook)

A chandelle is a maximum performance, 180° climbing turn that begins from an approximately straight-and-level flight and concludes with the airplane in a wings-level, nose-high attitude just above stall speed. Chandelle is a French word meaning "candle." French World War I pilots called the maneuver monter en chandelle, roughly translated as "to climb vertically," or "zoom."

The goal is to gain the most altitude possible for a given bank angle and power setting; however, the standard used to judge the maneuver is not the amount of altitude gained, but by **the pilot's proficiency as it pertains to maximizing climb performance for the power and bank selected, as well as the skill demonstrated.**



Chandelle maneuver

A chandelle is best described in two specific phases: the first 90° of turn and the second 90° of turn. The first 90° of turn is described as constant bank and changing pitch; and the second 90° as constant pitch and changing bank. During the first 90°, the pilot will set the bank angle, increase power and pitch at a rate so that maximum pitch-up is set at the completion of the first 90°. If the pitch is not correct, the airplane's airspeed is either above stall speed or the airplane may aerodynamically stall prior to the completion of the maneuver. Starting at the 90° point, the pilot begins a slow and coordinated constant rate rollout so as to have the wings level

Appendix A (CONTINUED)

when the airplane is at the 180° point while maintaining the constant pitch attitude set in the first 90°. If the rate of the rollout is too rapid or sluggish, the airplane either does not complete or exceeds the 180° turn as the wings come level to the horizon.

Prior to starting the chandelle, the flaps and landing gear (if retractable) should be in the UP position. The chandelle is initiated by properly clearing the airspace for air traffic and hazards. The maneuver should be entered from straight-and-level flight or a shallow dive at an airspeed recommended by the manufacturer—in most cases this is the airplane's design maneuvering speed. [Figure, A] After the appropriate entry airspeed has been established, the chandelle is started by smoothly entering a coordinated turn to the desired angle of bank; once the bank angle is established, which is generally 30°, a climbing turn should be started by smoothly applying elevator back pressure at a constant rate while simultaneously increasing engine power to the recommended setting. In airplanes with a fixed-pitch propeller, the throttle should be set so as to not exceed rotations per minute (rpm) limitations; in airplanes with constant-speed propellers, power may be set at the normal cruise or climb setting as appropriate. [Figure, B]

Since the airspeed is constantly decreasing throughout the chandelle, the effects of left-turning tendencies, such as P-factor, becomes more apparent. As airspeed decreases, right-rudder pressure is progressively increased to ensure that the airplane remains in coordinated flight. The pilot should maintain coordinated flight by sensing slipping or skidding pressures applied to the controls and by quick glances to the ball in the turn-and-slip or turn coordinator.

At the 90° point, the pilot should begin to smoothly roll out of the bank at a constant rate while maintaining the pitch attitude set in the first 90°. While the angle of bank is fixed during the first 90°, recall that as airspeed decreases, the overbanking tendency increases. [Figure, C] As a result, proper use of the ailerons allows the bank to remain at a fixed angle until rollout is begun at the start of the final 90°. As the rollout continues, the vertical component of lift increases; therefore, a slight release of elevator back pressure is required to keep the pitch attitude from increasing.

When the airspeed is slowest, near the completion of the chandelle, right rudder pressure is significant, especially when rolling out from a left chandelle due to left adverse yaw and left turning tendencies, such as P-factor. [Figure, D] When rolling out from a right chandelle, the yawing moment is to the right, which partially cancels some of the left turning tendency's effect. Depending on the airplane, either very little left rudder or a reduction in right rudder pressure is required during the rollout from a right chandelle. At the completion of 180° of turn, the wings should be leveled to the horizon, the airspeed should be just above stall speed, and the airplane's pitch high attitude should be held momentarily.

[Figure, E] Once demonstrated that the airplane is in controlled flight, the pitch attitude may be reduced and the airplane returned to straight-and-level cruise flight.

Appendix B

Demographic Survey

1. Which gender do you most closely identify with? (   )

a. Female b. Male c. Other d. Prefer not to answer
2. What is your age? (_____)
3. Do you hold a private pilot certificate (PPL)?
a. Yes b. No
4. Have you started Commercial pilot training?
a. Yes   b. No
5. Approximately how many years have you been flying?
(_____)
6. How many logged flight hours do you have?
(_____)
7. Approximate total number of hours in a Flight Training Device (FTD)?
(_____)
8. Approximate total number of hours in conventional PC-based flight simulation?
(_____)
9. Approximate total number of hours in Virtual reality flight simulation?
(_____)
10. Approximate total number of hours in other Virtual reality applications?
(_____)

11. You would like to try on a Virtual reality flight simulation.

1	2	3	4	5
strongly disagree	disagree	neutral	agree	strongly agree

11. You would like to use video instruction for learning flight maneuvers

1	2	3	4	5

12. You would like to use a conventional PC-based flight simulation for learning and practice.

1	2	3	4	5
strongly disagree	disagree	neutral	agree	strongly agree

13. You would like to use a Virtual reality flight simulation for learning and practice.

1	2	3	4	5
strongly disagree	disagree	neutral	agree	strongly agree

Appendix C

User Experience Survey

For the assessment of the flight simulation application, please fill out the following questionnaire. The questionnaire consists of pairs of contrasting attributes that may apply to the flight simulation. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression.

Example:

attractive	○	⊗	○	○	○	○	○	unattractive

This response would mean that you rate the application as more attractive than unattractive.

Please decide spontaneously. Don’t think too long about your decision to make sure that you convey your original impression.

It is your personal opinion that counts. Please remember: there is no wrong or right answer!

Please assess the product now by ticking one circle per line.

Appendix C (CONTINUED)

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Appendix C (CONTINUED)

Please rate your overall experience with the training method you are using for today's training activity.

1	2	3	4	5
Not satisfied	Slightly satisfied	Moderately satisfied	Very satisfied	Extremely satisfied

You would like to continue using this training method for future training and practice.

1	2	3	4	5
strongly disagree	disagree	neutral	agree	strongly agree

Appendix D

POST-TRAINING KNOWLEDGE TESTS.

1. What is the minimum altitude requirement to perform Chandelle?

A)1000 feet above ground level

B) 1500 feet above ground level

C) 2000 feet above ground level

2. Pilots who initiate a chandelle with a bank that is too steep will most likely

A) stall before completing the maneuver.

B) turn more than 180° before completing the rollout.

C) perform a comparatively level steep turn with a nose-high rollout at the 180° point

3. When performing a chandelle, where should maximum pitch occur?

A) 90° point.

B) 180° point.

C) 45° point.

4. Which best describes pitch and bank during the first half of the chandelle?

A) Changing pitch and bank.

B) Constant bank and changing pitch.

C) Constant pitch and bank.

5. Which best describes pitch and bank during the second half of the chandelle?

A) Constant bank and changing pitch.

B) Changing pitch and bank.

C) Constant pitch and changing bank.

Appendix D (CONTINUED)

6. What may occur if the initial bank is too shallow when performing a chandelle?

A) Stalling the aircraft before reaching the 180° point.

B) Completing the maneuver with too low a pitch attitude.

C) Completing the maneuver with excessive airspeed.

7. What the angle of bank should be when performing the first half of a chandelle?

A) 15°

B) 30°

C) 45°

8. What's the first step to perform a chandelle?

A) Clear the area.

B) Establish the appropriate entry configuration, power, and airspeed.

C) Select a prominent visual reference point

9. When performing a chandelle, where should you start the recovery at ?

A) 45° point

B) 90° point.

C) 180° point.

10. When performing a chandelle, once the bank angle is established, a climbing turn should be started by smoothly applying elevator back pressure at a constant rate while simultaneously____

A) increasing engine power

B) maintaining engine power

C) decreasing engine power

Appendix E

Self-efficacy Questionnaire

This questionnaire is designed to help us get a better understanding of how student pilots understand the concepts and tasks associated with the Chandelle maneuver. Please rate your degree of confidence by recording a number a 7-point scale (1 as not at all, 7 as very) of how confident you think you are at the following items.

	Rate (1 as not at all, 7 as very)
I understand the goal to perform a Chandelle.	
I understand the procedures to perform a Chandelle	
I feel confident of my ability to perform a Chandelle by myself.	