EMBRY-RIDDLE Aeronautical University
SCHOLARLY COMMONS

Journal of Aviation/Aerospace Education & Research

Volume 33 | Number 5 Article 2

2024

User Experience, Motivation and Self-Efficacy Comparisons Between Virtual Reality and Conventional PC-Based Flight Simulation Training

Tianxin Zhang Ph.D.
Tuskegee University, yokeymore@foxmail.com

Christina Frederick Ph.D.
Embry-Riddle Aeronautical University, christina.frederick@erau.edu

Andrew Dattel Ph.D.
Embry-Riddle Aeronautical University, andy.dattel@erau.edu

Elizabeth Blickensderfer Ph.D.
Embry-Riddle Aeronautical University, elizabeth.blickensderfer@erau.edu

Barbara Chaparro Ph.D.
Embry-Riddle Aeronautical University, chaparb1@erau.edu

Scholarly Commons Citation
Zhang, T., Frederick, C., Dattel, A., Blickensderfer, E., & Chaparro, B. (2024). User Experience, Motivation and Self-Efficacy Comparisons Between Virtual Reality and Conventional PC-Based Flight Simulation Training. *Journal of Aviation/Aerospace Education & Research, 33*(5). DOI: https://doi.org/10.58940/2329-258X.2092

JAAER

User Experience, Motivation and Self-Efficacy Comparisons Between Virtual Reality and Conventional PC-Based Flight Simulation Training

Tianxin Zhang[1a], **Christina Frederick**[2b], **Andrew Dattel**[2c], **Elizabeth Blickensderfer**[2d], **Barbara Chaparro**[2e]
[1]Tuskeegee University, AL 36088 USA
[2]Embry-Riddle Aeronautical University, FL 32124 USA,
[a]tzhang@tuskegee.edu, [b]christina.frederick@erau.edu, [c]andy.dattel@erau.edu, [d]elizabeth.blickensderfer@erau.edu
[e]chaparb1@erau.edu,

Abstract

The current project investigated the effect of utilizing Virtual Reality (VR) technologies for flight training by comparing user experience, motivation, and self-efficacy when using conventional desktop flight simulation versus VR flight simulation. This research employed a quasi-transfer of training experiment with 48 participants. Findings showed that VR flight simulation could provide a better user experience and generate a higher level of training motivation than traditional training technologies, while maintaining trainees' perceptions of self-efficacy. This work contributed positive evidence that VR flight simulation has great potential to be an effective means of flight training and provided a foundation for future research to continue exploring the training effects of VR flight simulation.

Keywords: *Virtual Reality, Flight Training, User Experience, Self-Efficacy, Motivation*

Introduction

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, Personal Computer-Based 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 PCATD's was generally positive and substantial when new tasks were introduced (Beckman, 2000, 2003; Homan & Williams, 1997; Taylor et al., 1999).

With the advent of Virtual Reality (VR) technologies, training is one of the most popular fields of VR application. With advantages like high-fidelity presentation that uses advanced graphics, detailed textures, and precise motion tracking, 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; Grabowski & Jankowski, 2015; Maytin et al., 2015; Vankipuram et al., 2010). 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 can be used effectively for pilot training, VR flight simulation has the potential to become the new low-cost alternative training method for novice pilots. However, currently, a limited number of studies have focused on evaluating the learning effectiveness of flight training utilizing VR technologies or the user experience of VR flight simulation. The effectiveness of flight training with VR simulation and student pilots' attitudes toward using VR for training are not clear.

Using VR simulation effectively for training requires the trainee to develop skills to use the technology, know the interface, and develop familiarity with the controls. Thus, 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 the user's willingness to adopt such technologies for training.

Purpose of Study

The current project investigated the effect of VR flight simulation versus conventional PCATD training on user motivation and satisfaction. Additionally, this project examined user experience resulting from training using VR flight simulation. While a primary focus of the project was the use of VR simulation for flight training, the ideas and findings generated by the current project could generalize to other areas within VR simulation-based training.

Literature Review

Simulation-Based Training 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 uses a constructivist approach, which means it utilizes 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., 2008a).

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 the Full Flight Simulator (FFS). FFS are multi-million-dollar machines with highly sensitive hydraulics and full visual displays. The second type of training device is the Flight Training Device (FTD). FTDs generally replicate an aircraft cockpit and often have a visual display system but provide no motion feedback. The third type of training device is the Aviation Training Device (ATD), and PC-based ATDs (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, and performance of the aircraft in terms of pitch, bank, and yaw (Beckman, 2000; Federal Aviation Administration, 2018).

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). It is natural to assume that the higher the level of fidelity, the greater 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., 2008b). 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. In contrast, Vaden and 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. Other experts have tried to explain the contradictory findings. For example, 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 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 skills to be learned and transferred (Liu et al., 2008b).

Regardless of simulation fidelity, a coherent structure has been proposed to understand training effectiveness. Kirkpatrick and Kirkpatrick (2006) articulated a four-level model for measuring training effectiveness. In this model, the first level is reaction, which reflects how the trainees perceive the effectiveness of training. The second level is learning, which describes 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 & Kirkpatrick, 2006). The fourth level is results, which are measured by trainees' achievement and implementation of the desired training goals over time to improve job performance and trainee morale.

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 & Dreyer, 2017; 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 recently (Robertson & Zelenko, 2014). As technologies advance, hardware and software have merged 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).

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 virtual environments can effectively simulate various conditions of work and life while successfully supporting learning processes (Norris et al., 2019; Sanchez-Vives & Slater, 2005). 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 generates the 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. VR flight simulation could be a valuable tool to improve the reaction level of training effectiveness in the Kirkpatrick Kirkpatrick (2006) model.

Many studies have been conducted to evaluate the effectiveness of training methods using VR technology. Sacks et al. (Sacks et al., 2013) tested the hypothesis that VR would be feasible and effective for construction safety training. 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 slide shows and a VR training group. The learning experience questionnaire results showed a significant advantage for VR over traditional training. Additionally, it was observed that the VR trainees were more focused than the other trainees. Overall, the researchers concluded that instruction using VR was more effective than safety training with traditional classroom presentations (Sacks et al., 2013).

Parong and Mayer (2018, 2021) conducted two experiments to evaluate the instructional effectiveness of immersive VR to teach scientific knowledge, as well as to examine the efficacy of adding a generative learning strategy to a VR lesson. In their first experiment, college students learned about how cells in the human bloodstream work using either immersive VR or a self-directed PowerPoint slideshow on a desktop computer. Each student completed a pre-questionnaire about their knowledge of the human body, a post-questionnaire about their experiences with the lesson, and a post-training knowledge test on the material they viewed during the lesson. The results showed that the slideshow group scored significantly better than the VR group on the post-test factual questions but not on the conceptual questions. However, the VR group rated the learning experience significantly higher in enjoyment, engagement, and motivation. In their second experiment, the slideshow was replaced with a segmented VR lesson with participants writing a summary after each segment. The procedure was the same and the segmented VR group performed better than the VR group for the factual and conceptual questions in the post-test (Parong & Mayer, 2021).

Cooper et al. (2021) studied how augmented multisensory cues affect performance in real environments. Multisensory cues are aspects of VR that provide input for multiple senses. The cues for this research included auditory (speakers for sounds), visual (goggles and color changing to demonstrate interaction with the simulation), and tactile (vibration) inputs. The study found that VR training with these augmented cues has a beneficial effect not only on performance and user experience in the virtual environment, but also on real task performance. Cooper et al. also reported that people using VR are more motivated to learn.

Overall, evidence from the limited research available indicates that VR is a promising training technology across domains. While VR training may not always enhance performance outcomes, the experience provided by VR training may positively impact user motivation.

VR Flight Simulation Related Research

Researchers are beginning to explore how VR technology can be used in flight training. Oberhauser and Dreyer (2017) conducted a series of experiments involving commercial airline pilots and non-pilots to evaluate the fidelity and usability of a VR simulation compared to a full flight simulator. The experimenter took the role of the co-pilot to ensure conformity with the flight tasks. Participants' heart rates and eye tracking heat maps were collected during the experiments. Results showed that users' overall operationals in VR are comparable to the full flight simulator environment; the VR flight had a sufficient level of simulation fidelity and a sufficient level of usability for pilots to fulfill tasks like flying the aircraft (Oberhauser & Dreyer, 2017). Oberhauser and Dreyer concluded that VR flight simulation could be a valuable tool to gather reliable information on human factors-related aspects of pilot behavior in flights.

Oberhauser et al. (2018) followed up with a training study that compared pilot performance using VR flight simulation versus conventional FTD using a within-subject design with experienced pilots. It was found that pilot workload demands measured by the NASA-Task Load Index (Hart & Staveland, 1988) were significantly higher in the VR condition than the FTD condition. In addition, participants' self-ratings of performance were also significantly lower in VR. Confounding factors were present in this study, including the participants' lack of familiarity with the VR interface, inaccuracies in the virtual hand model, the VR display latency to the control input, and the limited field of view from the head-mounted display (HMD) system (Oberhauser et al., 2018). Interestingly, the degradation of in-flight performance was not critical to safely conducting the given flight task in the VR environment. Even though the current VR flight simulation was found to have several disadvantages compared to the conventional flight simulation, it still exhibited potential to be a valuable tool for training and research purposes as advances in the field of VR technology occur (Oberhauser et al., 2018).

More recently, Fussell and Hight (2021) conducted usability testing of a VR flight training program compared with a conventional 2D simulation. Participants reported that they enjoyed using the 2D and VR simulations for training and found them useful, however, they rated the user experience attributes of the VR simulations higher than the 2D simulation overall. The researchers argued immersive features associated with VR may elicit a more positive experience.

Summary

The flight simulations available for pilots today include many improvements in terms of fidelity, usability, convenience, cost efficiency, and variety when compared to those of the past several decades (Kabashkin et al., 2023). With the advent of VR technology, the future development of flight simulations could present pilots with an even greater fidelity flight experience for training and practice.

For decades, researchers have found that flight simulation is an effective training tool, especially for concept and procedure learning (Taylor et al., 1999). More recently, researchers using newer VR technology showed beneficial training outcomes in several domains, including VR in medicine, high-hazard work, military training, and general science learning (Greunke & Sadagic, 2016). VR-based training is still relatively new in aviation training, and the effect of VR flight simulation training compared to traditional PCATD has not yet been fully studied and understood.

In studying VR-based training in flight environments, there are two overall questions of interest. The first question pertains to actual flight performance and whether VR-based training is as effective as more traditional methods. The second question addresses the perceived user experience of VR-based flight training. What is the user experience of VR-based flight training and do users experience self-efficacy and positive motivational experiences because of VR-based training? The current study examines this second question of interest. Lessons learned from existing research added valuable information to the development of the current study, which examined the user experience and motivation of VR flight simulation training, compared to desktop training and a control condition (Kirkpatrick & Kirkpatrick, 2006). For the current study, it is hypothesized that:

- H01: There is no difference in user experience across training methods.
- H02: Trainee motivation does not differ across training methods.
- H03: Trainees' self-efficacy pertaining to the trained flight maneuver will not differ across the training methods.
- H04: There is no difference in post-training self-efficacy on the selected flight maneuver among the 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 design in that a high-fidelity flight simulation rather than an aircraft is used to assess training tasks. Quasi-transfer of training has been used successfully in several flight simulation experiments (Taylor et al., 1993).

Participants

Advertisements for participants were distributed in two flight training locations in Florida. Forty-eight participants were recruited and were paid USD$30 for their participation. A priori power analysis indicated that 48 participants were sufficient to achieve adequate power. Table 1 below provides complete demographic information about participants. An ANOVA showed that there were no significant differences across the three groups in terms of Logged Flight Hours and FTD Hours. All participants had solo experience or a private pilot license but had not had commercial training.

Table 1

Demographic Information for VR, Desktop, and Control Groups

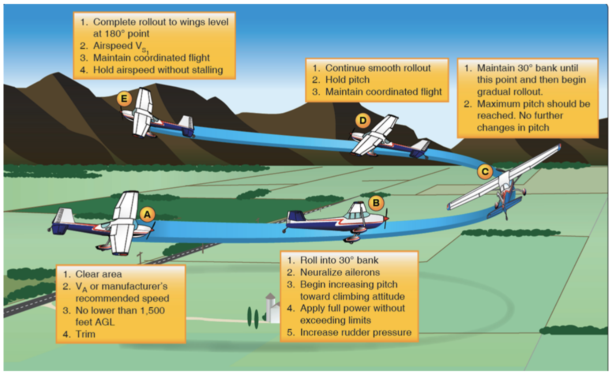
	VR Group	Desktop Group	Control Group
Number of Participants	18	16	14
Participant Age	M= 19.33, SD= 1.37	M= 20.31, SD=2.44	M=22.14, SD=5.71
Participant Gender	Male= 14, Female=4	Male=14, Female=2	Male=11, Female= 3
Logged Flight Hours	M= 111.01, SD=58.44	M=100.41, SD=32.75	M=118.96, SD=59.22
FTD Hours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

Apparatus and Materials

The training task that was chosen for this study was a flight maneuver from commercial pilot curriculums, the Chandelle. Figure 1 demonstrates the flight pattern of a Chandelle (Federal Aviation Administration, 2021).

Figure 1

Chandelle



Note. From Airplane Flying Handbook Chapter 10: Performance Maneuvers (Federal Aviation Administration, 2021)

Standardized written instructions for the Chandelle maneuver were given to all three groups before their training activities (Federal Aviation Administration, 2021). 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. A 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 (University of North Dakota, 2021).

The PCATD used for the Desktop group was a Windows-based personal computer running X-Plane 11(Version 11.51) (Laminar Research, 2017) with flight control accessories and three monitors. X-Plane 11 (https://www.x-plane.com/) 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 a 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 2 shows a participant using the VR flight simulation for practice.

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.

Figure 2

Study Participant Using the VR Flight Simulation

Note. Photograph by Authors

Measures

The measures included a demographic questionnaire, user experience surveys, a self-efficacy questionnaire, and a measure of motivation. The demographic questionnaire was completed before the first training activity, including age, gender, flight hours, and FTD hours. User experience was measured after the training activity was completed. The demographic questionnaire used in this study included questions about participants, such as age, gender, pilot rating, information about their flight experience specifically, total flight hours, years of flight experience, 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 surveys for user experience 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 has been used for a number of VR studies (Anton et al., 2018; Somrak et al., 2019; Su 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 the current project sample was .90 for Attractiveness, .77 for Perspicuity, .72 for Efficiency, .73 for Dependability, .79 for Stimulation, and .86 for Novelty.

The self-efficacy questionnaire was designed by adapting items from the General Self-Efficacy scale (Schwarzer & Jerusalem, 1995). The items on the 7-point self-efficacy questionnaire assessed participants' (1 - not at all, 7 - very) confidence levels about their understanding of the goal of the Chandelle, the procedures of the Chandelle, and their ability to perform a Chandelle. The purpose of using the self-efficacy questionnaire was to assess participants' affective outcomes resulting from training on reaction level. In this study, the Cronbach's alpha reliability for the self-efficacy questionnaire was .84.

The motivation items were developed for use in the present study. The items assessed preferences and satisfaction levels regarding different flight training methods. Participants were asked to rate their experiences with each method, ranging from overall satisfaction to the likelihood of continuing its use in future training sessions.

Procedure

Before the experimental session, the participants completed the demographic questionnaire online with questions about their willingness to use VR and desktop flight simulation for learning (pre-experiment).

Participants were then 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 training groups were a control group, a desktop group, and VR group. All three groups started training by reading a written explanation of the Chandelle maneuver (written learning, 10 minutes). Then, all three groups watched a video instruction of the Chandelle maneuver (15 minutes).

The first group was 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 two trials (20 minutes). After each trial, the participant could go back to review the written instruction.

The second group was 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 the VR group. Like 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 trials (20 minutes). After each trial, the participant could go back to review the written instruction.

All three groups were asked to complete the self-efficacy questionnaire right after reading the written instruction of the Chandelle maneuver and once again after the training practice. The user experience survey, and the motivation measure were also completed after the training practice.

Results Differences in User Experience Across Training Groups

This analysis assessed user experience across the training groups. The UEQ scores of VR and Desktop groups are shown in Table 2.

Table 2

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

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 Attractiveness (*M*=1.77, *SD*= .88) was significantly higher than the Desktop Group (*M*=1.03, *SD*=1.17), $t(32)=-2.09$, p=.04. Attractiveness 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 H01 was rejected.

Motivation Across Training Groups

Paired t-tests were run on the responses to the questions about participants' willingness to use VR and desktop flight simulation for learning (pre-experiment). An

ANOVA test was run on the response to questions about users' willingness to use the training method for future training from the UEQ. 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 the 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, H02 was rejected. Post Hoc test results are presented in Table 3.

Table 3

ANOVA Test Post Hoc Test Results for Hypothesis 2

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

Differences in Self-Efficacy Across Training Methods

This analysis examined self-efficacy across training groups. Time (pre-practice & post-practice), and the training technique (VR, Desktop, and Control) were set as independent variables, 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, H03 was rejected, and H04 was retained.

Discussion

User Experience, and Motivation: Differences Across Training Groups

Hypotheses 1 and 2 focused on user experience and motivation for VR flight simulation. The UEQ was used to measure user experience and customized questions were used to measure user motivation.

In a comparison of those in the VR flight simulation group score to UEQ benchmark scores (Schrepp et al., 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.

Current data suggests that the participants can perform their tasks quickly in a pragmatic way with the VR simulation, which could be as efficient as when using the desktop simulation. The participants generally liked to use the VR flight simulation and found it more attractive than using the conventional desktop simulation as the former is more innovative and creative.

The VR flight simulation group's stimulation score was 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 responses to questions about users' willingness to use the experimental training method for future training. Also, 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 received training in VR during the study. Those findings were consistent with prior research (Kwon, 2019; Makransky et al., 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 participants find the features of the VR flight simulation, such as the realistic 3D environment, 360 degrees of dynamic field of view, and the detailed interactive interface, beneficial to training, or that the VR interaction is closer to the real flight experience. Another factor that contributed to the results could be that participants were excited to use the new technology.

The findings of the present study correspond to findings from another VR study on user experience and motivation. Carbonell-Carrera et. al. (2021) concluded that VR use results in high scores in interest/enjoyment, perceived competence, effort/importance, pressure/tension, and value/usefulness on the Intrinsic Motivation Inventory (McAuley et al., 1989), which motivated the participants in their study to use VR.

Analyzing the current study results, we can see that participants generally like to use VR flight simulation, and their overall user experience is in some ways higher than using a desktop simulation. In other words, the VR reaction level of training effectiveness is higher than other techniques. It is important that the trainees enjoy the training and feel that it was a valuable experience (Thomas, 2018), which will motivate them to use the training technique again. With a potentially better user experience, VR flight simulation could be a valuable alternative to conventional desktop flight simulation as a training and learning tool.

Self-Efficacy and Training Results

Hypotheses 3 and 4 centered on the effect of different training methods on self-efficacy, and the correlation between self-efficacy and training results. Hypothesis 3 proposed that the participants' self-efficacy pertaining to the trained flight maneuver would not change after practice in all three groups. Hypothesis 4 proposed that there was no difference in post-training self-efficacy on the selected flight maneuver among the three groups. Results 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 their 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 an increase 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. These findings are consistent with Buttussi and 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.

Current findings support 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 a private pilot license (PPL) or were in the middle of PPL training (having passed their solo flight). Different results might be found if using pilots with various levels of experience or recruiting. The limited age demographic of the participants (M=20.48, SD=3.59) may also influence the results. 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. Lastly, participants may have felt undue pressure to provide positive feedback on the training session. These constraints may make the findings less generalizable than studies involving larger, more randomized samples.

To determine the true effects of VR flight simulation on pilot training results, future work is needed to examine the validity of the current findings with a long-term training intervention and 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. Additionally, it would be helpful to look at training effects for specific learning objectives, outcomes, and retention in future studies.

Conclusion

This study was conducted to compare VR flight training to traditional desktop training using PCATD. The findings are that the VR group had higher motivation and user experience. However, the results showed that there was no statistically significant difference in self-efficacy between VR, Desktop, and Control groups though there was an increase in self-efficacy after practice in all three groups. Overall, results support the viability of VR for ground training of student pilots due to higher student motivation and better user experience. Despite its limitations, this study shows the potential that VR has as a flight simulation tool. This potential will only grow as the technology of VR advances.

References

Alessi, S. M. (1988). Fidelity in the design of instructional simulations. *Journal of Computer-Based Instruction*, *15*(2), 40–47.

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

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (pp. 71–81, Vol. 4). Academic Press.

Beckman, W. (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). https://doi.org/10.15394/jaaer.2000.1244

Beckman, W. (2003). An 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

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

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

Carbonell-Carrera, C., Saorin, J. L., & Diaz, D. M. (2021). User VR experience and motivation study in an immersive 3D geovisualization environment using a game engine for landscape design teaching. *Land*, *10*(5), 492. https://doi.org/10.3390/land10050492

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

Cooper, N., Millela, F., Cant, I., White, M. D., & Meyer, G. (2021). Transfer of training-virtual reality training with augmented multisensory cues improves user experience during training and task performance in the real world. *PloS One*, *16*(3). https://doi.org/10.1371/journal.pone.0248225

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

Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, *26*, 43–71. https://doi.org/10.1002/piq.21143

Federal Aviation Administration. (2018). *FAA approval of aviation training devices and their use for training and experience (Advisory Circular 61-136B)* (tech. rep.). U.S. Department of Transportation. https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentid/1034348

Federal Aviation Administration. (2021). *Chapter 10: Performance maneuvers* (tech. rep.). U.S. Department of Transportation. https://www.faa.gov/sites/faa.gov/files/regulations_policies/handbooks_manuals/aviation/airplane_handbook/11_afh_ch10.pdf

Fussell, S. G., & Hight, M. P. (2021). Usability testing of a VR flight training program. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 1124–1128. https://doi.org/10.1177/1071181321651096

Geršak, G., Lu, H., & Guna, J. (2018). Effect of VR technology matureness on VR sickness. *Multimedia Tools and Applications*, *79*, 14491–14507. https://doi.org/10.1007/s11042-018-6969-2

Grabowski, A., & Jankowski, J. (2015). Virtual reality-based pilot training for underground coal miners. *Safety Science*, *72*, 310–314. https://doi.org/10.1016/j.ssci.2014.09.017

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

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in psychology* (pp. 139–183, Vol. 52). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

Homan, W. J., & Williams, K. D. (1997). Effects of PC-based pretraining on pilots' performance in an approved flight-training device. *Journal of Avi-*

ation/Aerospace Education & Research, *7*(3). https://doi.org/10.15394/jaaer.1997.1198

Kabashkin, I., Misnevs, B., & Zervina, O. (2023). Artificial intelligence in aviation: New professionals for new technologies. *Applied Sciences*, *13*(21). https://doi.org/10.3390/app132111660

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd). Berrett-Koehler.

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

Laminar Research. (2017). X-plane 11.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger (Ed.), *Hci and usability for education and work* (pp. 63–76). Springer. https://doi.org/10.1007/978-3-540-89350-9_6

Liu, D., Blickensderfer, E., Vicenzi, D., & Macchiarella, D. (2008a). Transfer of training. In D. Vincenzi & J. A. Wise (Eds.), *Human factors and simulation* (pp. 35–50). CRC Press, Taylor & Francis Group.

Liu, D., Vicenzi, D., & Macchiarella, D. (2008b). Simulation fidelity. In D. Vincenzi & J. A. Wise (Eds.), *Human factors and simulation* (pp. 51–66). 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

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58. https://doi.org/10.1080/02701367.1989.10607413

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] [Unpublished doctoral dissertation].

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

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

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *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, May). Voices from the past: An oral history of a technology whose time has come again. http://www.theverge.com/a/virtual-reality/oral_history

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., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339. https://doi.org/10.1038/nrn1651

Schwarzer, R., & Jerusalem, M. (1995). General self-efficacy scale. https://doi.org/10.1037/t00393-000

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

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*, *24*, 241–254. https://doi.org/10.1007/s10055-019-00394-w

Talleur, D. A., Taylor, H. L., Emanuel, T. W., Rantanen, E., & Bradshaw, G. L. (2003). Personal computer aviation training devices: Their effectiveness for maintaining instrument currency. *The International Journal of Aviation Psychology*, *13*(4), 387–401. https://doi.org/10.1207/S15327108IJAP1304_04

Taylor, H. L., Lintern, G., Hulin, C. L., Talleur, D. A., Emanuel, 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. https://doi.org/10.1207/s15327108ijap0904_1

Taylor, H. L., Lintern, G., & Koonce, J. M. (1993). Quasi-transfer as a predictor of transfer from simulator to airplane. *The Journal of General Psychology*, *120*(3), 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] [Doctoral dissertation]. https://commons.erau.edu/edt/388

Thompson, N. T., Carroll, M. B., & Deaton, J. E. (2008). Justification for use of simulation. In P. A. Hancock, D. A. Vincenzi, J. A. Wise, & M. Mouloua (Eds.), *Human factors and simulation* (pp. 25–34). CRC Press, Taylor & Francis Group. https://doi.org/10.1201/9781420072846

University of North Dakota. (2021). Commercial steep turns [video] [March 2]. 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. https://doi.org/10.1207/s15327108ijap1504_5

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