Summer 8-2024

Technology Acceptance of Virtual Reality for Aircraft Maintenance Training

Lusine Carlsson
Embry-Riddle Aeronautical University, carlssol@my.erau.edu

Technology Acceptance of Virtual Reality for Aircraft Maintenance Training

Lusine Carlsson

Dissertation Submitted to the College of Aviation in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Aviation

Embry-Riddle Aeronautical University

Daytona Beach, Florida

August 2024

Technology Acceptance of Virtual Reality for Aircraft Maintenance Training

By

Lusine Carlsson

This dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, Dr. Steven Hampton , and has been approved by the members of the dissertation committee. It was submitted to the College of Aviation and was accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Aviation.

Steven Hampton Digitally signed by Steven Hampton
Date: 2024.08.13 14:50:18 -04'00'

Steven Hampton, Ed.D.
Committee Chair

Jennifer E. Thropp Digitally signed by Jennifer E. Thropp
Date: 2024.08.13 21:15:40 -04'00'

Jennifer Thropp, Ph.D.
Committee Member

Steven Hampton Digitally signed by Steven Hampton
Date: 2024.08.14 16:25:07 -04'00'

Steven Hampton, Ed.D.
Associate Dean, School of Graduate Studies, College of Aviation

Daniel Friedenzohn Digitally signed by Daniel Friedenzohn
Date: 2024.08.14 08:13:07 -04'00'

Daniel Friedenzohn, J.D.
Committee Member

Alan J. Stolzer Digitally signed by Alan J. Stolzer
Date: 2024.08.15 08:24:59 -04'00'

Alan J. Stolzer, Ph.D.
Dean, College of Aviation

Committee Member

Rachelle Strong Digitally signed by Rachelle Strong
Date: 2024.08.14 12:46:26 -07'00'

Rachelle Strong, Ph.D.
Committee Member (External)

July 15, 2024

Signature Page Date

Abstract

Researcher: Lusine Carlsson

Title: Technology Acceptance of Virtual Reality for Aircraft Maintenance Training

Institution: Embry-Riddle Aeronautical University

Degree: Doctor of Philosophy in Aviation

Year: 2024

A study was conducted to investigate Virtual Reality (VR) technology acceptance for aircraft maintenance training by student mechanics. The study employed a survey methodology using the extended Technology Acceptance Model (TAM) as the theoretical framework. Participants were student mechanics from Federal Aviation Administration (FAA) approved aircraft maintenance schools within the United States. The study attempted to address how the variables of *perceived usefulness (PU)* and *perceived ease of use (PEU)* influence the students' *behavioral intention (BI)* to use VR technology for aircraft maintenance training, and how external variables of *self-efficacy (SE)*, *perceived enjoyment (PE)*, *perceived health risk (PHR)*, *performance expectancy (PEXP)*, and *perceived behavioral control (PBC)* influence *perceived usefulness (PU)* and *perceived ease of use (PEU)* of VR technology.

Variables were measured through a survey instrument, utilizing a five-point Likert scale to collect quantitative data. The survey was administered digitally. First, a pilot study was conducted for the purposes of improving the survey instrument and the hypothesized model, followed by the main study. Confirmatory Factor Analysis (CFA) was applied to the pilot study data to investigate relationships between construct variables

and their observed variables and to confirm the theoretical model. A good model fit was not achieved with a small sample size of N=55. To encourage student participation, the final study scope was reduced to only include the TAM variables of PU, PEU, and BI. The final study produced 65 additional responses, for a total of 120 (including the pilot study data), which was deemed inadequate for confirmatory analysis. Instead, the responses were analyzed through the Exploratory Factor Analysis (EFA) technique to explore the data and extract a factor structure based on the measurement items (survey questions), and to address the research question,: What are the underlying factors of the student aircraft mechanic survey? The Principal Axis Factoring (PAX) with oblique rotation was selected for an extraction method. The EFA results produced two viable factors. Factor 1 included all measurement items that were thought to represent the constructs of BI and PU. It also included one measurement item from PEU. Factor 2 only included measurement items from PEU. The EFA results suggested that participants may have answered the questions on behavioral intention to use VR based on how useful they perceived VR to be. Therefore, Factor 1 was considered to be more representative of the perceived usefulness (PU) construct and Factor 2 perceived ease of use (PEU). The findings contribute to aviation literature with recommendations for the target population, research methodology, and future research. The study identified factors that could be considered in future confirmatory research to gain more insight on aircraft mechanic students' opinions on VR to make predictions on actual use of the technology for aircraft maintenance training. From a practical perspective, the findings suggest that schools and regulatory agencies should consider implementing VR for aircraft maintenance training in such a way that will make VR useful to students for gaining specific knowledge/skills without compromising its ease of use, and by taking steps to make VR an enjoyable experience for students.

Keywords: virtual reality (VR), perceived usefulness (PU), perceived ease of use (PEU), behavioral intention (BI), exploratory factor analysis (EFA), principal axis factoring (PAX), oblique rotation, direct oblimin

Dedication

This dissertation is dedicated to my family members:

To my two sons, David, and Daniel, who gave me unlimited hugs and kisses to keep me going during the most stressful times, worked as a team when I needed help to set up my workstation for the qualifying exam, and made sure I didn't forget to eat and drink during busy times. David, every time you walked up to me while I was working on my dissertation and said, "I'm so proud of you, mama," my stress disappeared, my motivation to continue working hard increased unbelievably, and I was reminded that failing is not an option. Daniel, your love, and caring means the world to me. You frequently brought me water and reminded me to drink it, because you knew I often did not have time to take care of myself. You helped me proofread my work and corrected grammar. I love you both, my sons. Thank you for being you!

To my husband, Chris, who started operating around my busy schedule from the moment I got accepted into the Ph.D. in Aviation program and did not complain, who tolerated my commands, made the most delicious breakfasts, and did everything possible for my comfort during long hours of work. Thank you for your love and caring!

Acknowledgments

This research was completed with the support of my Committee Chair, Dr. Steven Hampton, and Committee Members, Dr. Jennifer Thropp, Dr. Daniel Friedenzohn, and Dr. Rachelle Strong. I greatly appreciate my Committee Members in conducting reviews and providing valuable feedback to help me improve my research skills and the quality of the study. When I needed help with administering the survey online and recruiting participants through social media, Dr. Rachelle Strong helped me figure out how social media works and shared my survey post with her professional community. I am especially thankful to Dr. Hampton who took the initiative to coordinate with my Committee Members and with Dr. Dothang Truong when I needed help with study decisions. Dr. Hampton worked with me to find a feasible path forward to help me succeed when my study failed to achieve the required sample size. He also facilitated the data collection for my pilot study by getting me in touch with the Embry-Riddle Aeronautical University (ERAU) Aviation Maintenance Science Department and helping me obtain their support. I greatly appreciated Dr. Hampton's continuous caring, understanding of my personal situation, and words of kindness and encouragement, which kept me going during difficult times. I am also grateful to Dr. Dothang Truong for his time and caring answering questions related to the study's statistical analysis. His perceptive and respectful nature helped me build trust, feel confident and supported. Finally, I would like to thank the ERAU Aviation Maintenance Science Department faculty and their students for helping with the pilot study.

Table of Contents

Page

Abstract iv

Dedication vii

Acknowledgments viii

List of Tables xii

List of Figures xiii

Chapter I: Introduction 1

Statement of Problem 6

Purpose Statement 8

Research Questions and Hypotheses 8

Significance of the Study 11

Delimitations 12

Limitations and Assumptions 12

Summary 15

Definition of Terms 16

List of Acronyms 18

Chapter II: Review of the Relevant Literature 21

Virtual Reality Applications and Benefits 21

Background 21

VR Applied to Training and Education 24

VR Benefits and Implications on Aircraft Maintenance Training 27

Current State of VR & Technology Adoption 30

Theoretical Framework of the Study 34

Technology Acceptance Model 34

Relevant Research and Major Findings 38

Summary 42

Chapter III: Methodology 45

Research Approach 45

Design and Procedures 46

Apparatus and Materials 49

Population/Sample 50

Sources of the Data 54

Data Collection Device 55

Reliability and Validity 58

Treatment of the Data 60

Descriptive Statistics 60

Hypothesis Testing 62

Summary 63

Chapter IV: Results 65

Pilot Study Results 65

Final Study Results 71

Demographics 71

Descriptive Statistics 73

EFA Assumptions 76

EFA Analysis Results 77

Reliability and Validity 84

Summary 87

Chapter V: Discussion, Conclusions, and Recommendations 89

Discussion 90

Characteristics of Participants 90

Discussion of the EFA Decisions & Results 95

Conclusions 103

Theoretical Contributions 104

Practical Contributions 106

Limitations of the Findings 108

Recommendations for Future Research 110

Recommendations for the Target Population 111

Recommendations for Future Research Methodology 112

Recommendations for Future Areas of Research 115

References 117

Appendix 124

List of Tables

Table Page

1 Guidance for Minimum Sample Size Based on Model Complexity........................ 52

2 Model Fit Summary from SEM 62

3 Factor Loadings, Construct Reliability, Convergent & Discriminant Validity 70

4 Demographic Attributes of Participants.. 72

5 Final Survey Constructs, Observed Variables, & Survey Questions....................... 73

6 Descriptive Statistics Results 74

7 Construct Reliability Test Results: Cronbach's Alpha... 85

8 Final Study Sample Diversity Based on Volunteered Data 94

List of Figures

Figure Page

1 Hypothesized Relationships Between the Variables 10

2 The Relationship Between Stimulus-Organism-Response 35

3 Early Modified Version of TAM 36

4 Final 1996, TAM Model as Applied in this Study 37

5 Confirmed Hypotheses from Wang, Anne, and Ropp (2016)................................. 42

6 Preliminary Survey Questions 57

7 First Specified CFA Model 67

8 Model Fit Indices Results 68

9 Initial EFA Analysis Showing the Variance by a Single Factor 78

10 Initial EFA Analysis: Scree Plot Output 79

11 Final EFA Analysis Showing Variance 80

12 Final EFA Analysis: Pattern Matrix.. 81

13 Final EFA Analysis: Structure Matrix 82

14 Final EFA Analysis: Factor-to-Factor Correlation Matrix 84

15 Construct Reliability Test Results: Inter-Item Correlations................................... 86

16 Construct Reliability Test Results: Corrected Item-Total Correlations 86

17 Proposed Model Structure Based on EFA Results.. 105

Chapter I: Introduction

Aircraft maintenance practices have a direct impact on the safety of flight. The design of aircraft equipment, environmental conditions such as accessibility, temperature, lighting, mechanics' skills, and knowledge are all factors that can contribute to errors in maintenance and workplace safety risks (Federal Aviation Administration [FAA], 2011). Maintenance training typically addresses these factors through traditional teaching methods that emphasize basic concepts in physics and specific aircraft systems (e.g., hydraulic power systems, cabin atmosphere control systems, aircraft instruments, fuel system, engines, etc.). To become an aircraft mechanic, one must obtain an Aviation Maintenance Technician (AMT) certificate (with Airframe or Powerplant rating or both) through an FAA approved school. Additional training is required for new equipment introduced into service. The current requirements of the maintenance training/curriculum for an AMT certificate, as well as for airline maintenance, are defined by the FAA. Specifically, for the AMT certificate, the requirements are outlined in 14 C.F.R. Part 147, originally promulgated in 1968 and last updated in 2022 (Aviation Maintenance Technician Schools, 2021). The latest FAA rule shows progress towards competency-based training (Aviation Maintenance Technician Schools, 2022), which presents opportunities for AMT schools to advance their methods for delivering training. Still, it has been recognized, within the aviation industry, that the FAA requirements for maintenance training have fallen behind because the needs of aircraft mechanic training and education are very different today than that of decades ago due to technological advancements over time (White et al., 2000). It includes not just the training curriculum/content, but also the FAA approved methods for delivering the curriculum

(Goldsby & Soulis, 2002). The FAA rule requires AMT schools to show how their curriculum and methods will ensure building sufficient knowledge and skills in students, but it does not prescribe any specific, modern, and potentially more effective training tools for schools to adopt (Aviation Maintenance Technician Schools, 2021).

While the aviation industry is currently safer than it has ever been, maintenance related accidents and incidents still occur. For example, maintenance-related accident data for general aviation suggests no change in the rate of accidents in the time period of 1989-2013. The rate remained at 4.3 accidents per million flight hours. Improper maintenance and errors, such as mis-installations/improper reassembly, improper rigging, inadequate inspection, were found to be primary contributing factors to these accidents/incidents (Boyd & Stolzer, 2015). Historically, aircraft accident and incident cases related to improper maintenance are many. For example, in 2000, an incorrect installation of a spacer in an Air France aircraft's landing gear led to a tire blowout, loss of engines, fire, and the death of over 100 passengers and crew (International Aviation Safety Assessment [IASA], n.d.). In the same year, a DC-8-71F crashed on approach due to an improperly installed elevator control. All three crew members died (IASA, n.d.). The 2003 International Air Transportation Association (IATA) Safety Report findings also confirm the importance of proper aircraft maintenance. The data showed that 26% of accidents included maintenance related events as a contributing factor (Rankin, 2007). The report emphasizes the implications that maintenance errors have on airline operational costs. At a minimum, the industry data also suggests deficiencies in aircraft maintenance which could partially be addressed though more effective training methods than what we have today (Rankin, 2007).

The FAA instructional equipment requirements imposed on AMT schools do not ensure maximum exposure to various aircraft types and aircraft systems. Schools are required to conduct practical projects and demonstrations on an actual aircraft (any type) certified by the FAA for either private or commercial use, but it does not necessarily have to be in airworthy condition (Aviation Maintenance Technician Schools, 2021). Newly graduated aircraft mechanics are likely to be unfamiliar with all the aircrafts and systems that they will be maintaining during their careers, which could lead to mistakes and wrong maintenance decisions. Yet, it would not be feasible to require schools to provide practice on all types of aircraft and for technologies that vary in complexity. Some level of familiarity could be gained from static instructional materials or through videos, but such methods do not support hands-on/skill-building activities and therefore, may not be effective. The aviation maintenance industry can benefit from innovative training techniques that include modern technologies currently being utilized for similar purposes in other areas of practice which have been shown to be highly effective.

Virtual Reality (VR) technology is currently being studied and used for training within industries that require a high level of standardization and skill-based learning, including medical procedures, manufacturing operations, and pilot training (Goldenstein, 2020). VR offers an innovative way of implementing modern technology to enhance training. Instructional theories from a constructivist view suggest that real-life activities provide the context needed to improve learners' performance and motivation (Huang & Liaw, 2018). Constructivist theories suggest that training is more effective when students are engaged and active during the training rather than passive (SUNY: University at Buffalo, 2021). Constructivism also suggests that the social aspect is an inherent part of the learning process because it involves instructors and students working together to build knowledge and skills, which is another reason students need to be kept engaged. Finally, emphasis is placed on providing experiences that reinforce knowledge and facilitate learning (SUNY: University at Buffalo, 2021). VR technology can simulate real-life activities for training purposes, and unlike traditional teaching methods (e.g., classroom lectures, textbook), it helps students to be more engaged by providing the opportunity to practice tasks, interact with aircraft systems in a realistic maintenance environment, interact with the teacher and other students (in a multi-user mode), and actively utilize the information learned (Goldenstein, 2020). In basic terms, VR is a 3-dimensional (3D) simulation of an environment with the use of electronic equipment and computerized modeling. At a minimum, it includes a head mounted display (HMD), controls, and sensors mounted/attached/held by the user. Augmented Reality (AR) and Mixed Reality (MR) are similar to VR technologies, in that they also can create virtual environments through the use of computer devices. AR and MR provide an overlay of computer generated (CG) content on real-world objects, with the only difference being that with MR the CG content can exhibit occlusion because it is anchored to the real-world object. VR is the only technology that is able to provide a fully immersive experience within an environment that has been created using real-world or CG (synthetic) content or both (Irvine, 2017).

Compared to traditional training methods, VR simulation training not only reduces training costs, but also increases the effectiveness of training by modeling a realistic environment and kinesthetic movements and enabling significantly higher levels of interaction in team-based settings (Stone et al., n.d.). VR has also been studied for effectiveness in automobile maintenance (Borsci et al., 2015) and manufacturing operations training (Al-Ahmari et al., 2016), demonstrating benefits in error reduction and benefits associated with assembly decisions and training. VR provides a safer way for student aircraft mechanics to gain hands-on practice as there is no potential for damaging aircraft hardware in case of an error, and students can practice maintenance procedures as many times as needed. The risk of injury, especially during complex tasks, is also reduced due to the environment being virtual, where it lacks common environmental hazards anticipated during aircraft maintenance (e.g., slippery surfaces, fluid spillage, chemical fumes, working on a ladder or at unsafe heights, etc.). VR also enables working on different aircraft types to build familiarity. Aircraft maintenance schools are typically equipped with a single aircraft type for students to practice on as the bare minimum imposed by the FAA. With VR, various aircraft types become easily accessible as they are simulated. There is no need for the schools to purchase/store/maintain an actual aircraft. Finally, VR can be set up to allow multiple students to practice at the same time, because aircraft equipment availability would no longer be an issue. Many benefits of VR are transferable to aircraft maintenance training and well substantiated through the current literature. However, the user perceptions of VR and the question, "Would the aircraft mechanic students actually use the technology if made available?," has not yet been addressed.

This study investigated the use of Virtual Reality (VR) technology for aircraft maintenance training as it pertains to student mechanics' acceptance of the technology. The study was confirmatory and employed a theoretical framework, the extended Technology Acceptance Model (TAM), to help predict actual VR use by aircraft mechanic students based on how the technology is perceived and external factors that may influence those perceptions. TAM offers an opportunity to investigate the strength of these relationships as contributing factors of behavioral intention through a survey methodology.

Statement of Problem

Acceptance of VR could potentially lead to improvement of AMT training and FAA requirements by demonstrating that the technology is a viable option for providing better exposure and access to different aircraft types and systems compared to traditional training methods. VR technology is currently accepted and applied in a number of fields that have similar safety and security measures as aviation; however, within the aviation industry, its potential is only recently starting to be recognized. Regarding VR implementation in training, many studies have focused on effectiveness of the technology rather than acceptance. There is a significant gap in the literature addressing acceptance of VR technology, which is even larger for acceptance of the technology specifically for aviation maintenance training. Much of the literature on VR acceptance focuses on the use of the technology for medical training, design evaluation, manufacturing/assembly, and as a tool that could deliver more meaningful learning in general as compared to conventional learning methods. It is reasonable to anticipate an increase in interest in VR technology within the aviation industry, as ongoing research demonstrates its benefits in terms of facilitating learning and as a feasible, cheaper, and lower risk option. The current literature on VR for aircraft maintenance training substantiates the effectiveness of the technology but not acceptance of the technology by primary user groups.

The aviation industry recognizes potential cultural hurdles related to the implementation of VR technologies for training (SAE, 2016). Organizational challenges with VR use were observed in aerospace manufacturing and included security and configuration control concerns. These challenges have the potential to impose limitations on technology implementation (SAE, 2016), where user perception of usefulness/ease of use (factors that are thought to impact technology acceptance) could be impacted. It is hard to foresee exactly what challenges will arise with VR use for AMT training. Understanding factors that influence user perception and adoption of the technology could help with decisions on how to implement VR as a new baseline for training and how to overcome some of the challenges at an organizational level. The literature specific to VR technology adoption by aircraft mechanics for the purposes of maintenance training needs to be addressed in order to understand how the technology is currently being perceived by primary users and if aircraft mechanics would actually use the technology if implemented in the near future. One study that comes closest to this topic by Rupasinghe et al. (2011) shows promising results in terms of aircraft mechanic students providing positive feedback after completing VR training scenarios, but the study lacks a theoretical framework for making conclusions on acceptance of the technology by the students and behavioral intent to use the technology. A study by Fussell (2020) investigated VR technology acceptance using TAM and some of the same external variables as in this study but for flight training. The literature on VR for aircraft maintenance training using TAM is non-existent.

Purpose Statement

The purpose of this study was to use the extended TAM to investigate future aircraft mechanics' acceptance of VR technologies for their training. The model suggests that student mechanics' behavioral intention to use the technology predicts actual use, driven by factors of perceived usefulness and perceived ease of use (Lala, 2014). The study also investigated how perceived usefulness and perceived ease of use are influenced by external factors of self-efficacy (SE), perceived enjoyment (PE), perceived health risk (PHR), performance expectancy (PEXP), and perceived behavioral control (PBC).

Research Questions and Hypotheses

The research questions and hypotheses were generated based on the final 1996 version of TAM (Chuttur, 2009), as extended with the selected external variables based on the literature review.

RQ1

To what extent does *perceived usefulness* (PU) affect the student mechanics' intent to use VR technology for maintenance training? [Addressed through H1]

RQ2

To what extent does *perceived ease of use (PEU)* affect student mechanics' intent to use VR technology for maintenance training? [Addressed through H2.]

RQ3

What is the relationship between *perceived ease of use (PEU)* and *perceived usefulness (PU)*? [Addressed through H3]

RQ4

How do external variables of *self-efficacy (SE)*, *perceived enjoyment (PE)*, *perceived health risk (PHR)*, *performance expectancy (PEXP)*, and *perceived behavioral control (PBC)* influence *perceived usefulness (PU)* and *perceived ease of use (PEU)* of VR technology for maintenance training? [Addressed through H4 - H11]

The hypotheses are listed below and also shown in Figure 1:

H1

Perceived usefulness (PU) has a positive effect on student mechanics' *behavioral intention (BI)* to use VR technologies for maintenance training.

H2

Perceived ease of use (PEU) has a positive effect on student mechanics' *behavioral intention (BI)* to use VR technologies for maintenance training.

H3

Perceived ease of use (PEU) has a positive effect on *perceived usefulness (PU)*.

H4

Self-efficacy (SE) has a positive effect on *perceived ease of use (PEU)*.

H5

Self-efficacy (SE) has a positive effect on *perceived usefulness (PU)*.

H6

Perceived enjoyment (PE) has a positive effect on *perceived ease of use (PEU)*.

H7

Perceived enjoyment (PE) has a positive effect on *perceived usefulness (PU)*.

H8

Perceived health risk (PHR) has a negative effect on *perceived ease of use (PEU)*.

H9

Perceived health risk (PHR) has a negative effect on *perceived usefulness (PU)*.

H10

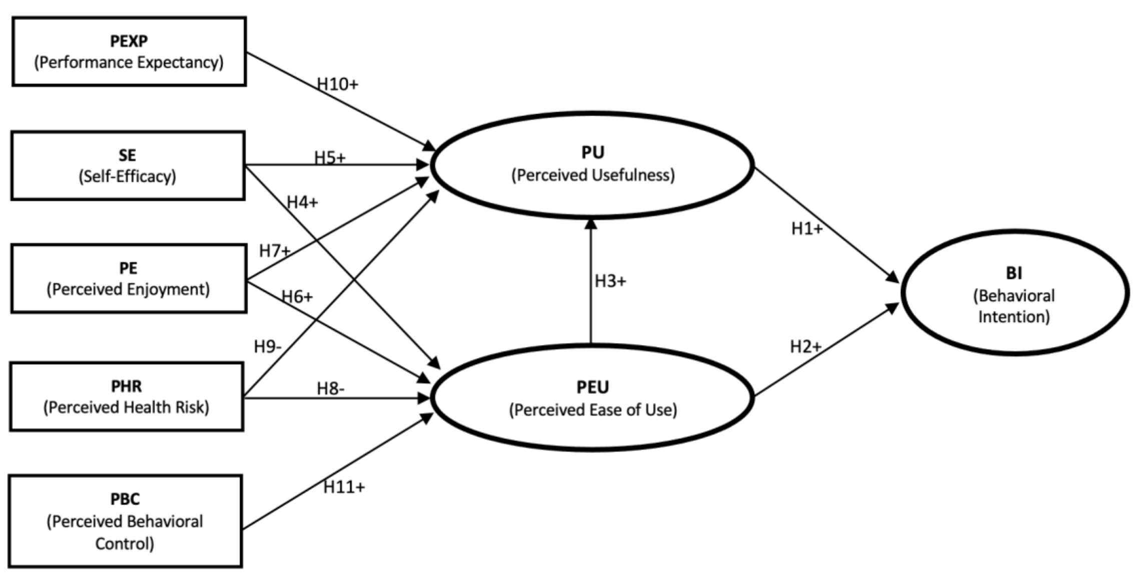
Performance expectancy (PEXP) has a positive effect on *perceived usefulness (PU)*.

H11

Perceived behavioral control (PBC) has a positive effect on *perceived ease of use (PEU)*.

Figure 1

Hypothesized Relationships Between the Variables



Significance of the Study

The study has theoretical significance in that it contributes to the literature within the aviation industry on VR technology adoption for maintenance training. The selected methodology enabled analysis based on a theoretical framework, using the TAM, to study student mechanics' acceptance of VR technology based on perceived usefulness and perceived ease-of-use as contributing factors to behavioral intention to use the technology. The model has been applied in many previous studies and deemed to be highly valid for making inferences about actual system use (Chuttur, 2009). TAM allows expansion of the model to account for external variables that may be associated with perceived usefulness and perceived ease-of-use. Upon reviewing relevant literature, there was no evidence of previously conducted research on this specific topic using TAM.

The study was developed to reveal how students in an aircraft maintenance program perceive VR technology, how their perceptions are impacted by external variables, and how their perceptions translate into intentions toward using such technologies (Park, 2009). The practical significance of this research is in providing additional information that will help airline operators, aircraft manufacturers, and aircraft maintenance training schools to assess the viability of adopting VR technology for aircraft maintenance training. The study could provide a better understanding of specific areas of concern or factors that impact future mechanics' perceptions/intentions that may potentially lead to resistance to using the technology. The results could be used to generate mitigation strategies/techniques for these specific areas of concern or serve as guidance for the development of business strategies related to instructional design using VR technologies.

Delimitations

A survey methodology was employed to collect data on student mechanics' acceptance of VR technology for aircraft maintenance training, using extended TAM (1996 final version) as the theoretical framework (Chuttur, 2009). The external variables of choice and hypothesized relationships were based on the literature review. Structured survey questions about candidate mechanics' perception of VR for aircraft maintenance training provided subjective and quantitative data that was analyzed to make inferences about intent to use the technology and predictions about actual use. The survey instrument was adapted from similar published studies, where TAM was used to investigate VR/AR technology acceptance within various contexts. The preliminary survey questions are presented in Figure 6, along with references to the study source. The questions have been customized to fit the context of the study and were further adjusted based on the results from the pilot study and feedback provided by subject matter experts. This tailoring approach was necessary to ensure validity of the survey instrument. To be cost-effective and as a more practical approach, the survey was solicited via email and social media to student mechanics enrolled in FAA approved schools within the U.S. The selected sampling method was judgement or convenience sampling -- a non-probabilistic strategy. The scope of the study was confirmatory, and as such, Structural Equation Modeling (SEM) was used to analyze the survey data and validate the model.

Limitations and Assumptions

External validity refers to the representativeness of the sample, making sure the results and conclusions are generalizable to the target population (Vogt et al., 2014). There was a risk of the online survey approach and the non-probabilistic/convenience sampling method compromising the external validity of the study by producing a smaller sample size than desired or a sample that is not diverse or comparable to the population demographics. Not all FAA maintenance school students were solicited to take the survey. Participation was limited by each school's willingness to support the study. The responses were anonymous and voluntary. Also, the data collected was subjective. This means there was some potential for bias and errors, though it is assumed that the students provided honest and accurate responses. To encourage participation, students were offered an incentive - an entry for a raffle to win a common aircraft maintenance tool (a torque wrench, valued up to $400) for fully completing the questionnaire (the demographic and Likert scale sections of the survey). Missing data in participants' responses would also present limitations if it was to the extent of having to delete the responses and leading to the success criteria not being met. Deletion of responses was only considered if the option of replacing the missing data with an average value was not viable.

The scope of the study was limited to candidate aircraft mechanics from 14 C.F.R. Part 147 - Aviation Maintenance Technician Schools. Therefore, generalizability was considered for the greater student mechanic population within the U.S. and not the general population of current aircraft mechanics. While 14 C.F.R. Part 65 addresses the knowledge and experience requirements of non-flight crewmembers for certification purposes, including ratings for Airframe and Powerplant (A&P) for mechanics, Part 147 is the control authority over AMT schools, in terms of curriculum, practical training, and equipment requirements; and therefore, Part 147 is considered more relevant to the scope of this study than Part 65. The study was limited to participants enrolled in an FAA approved AMT school as controlled by Part 147, regardless of the students' qualification status for Part 65 certification testing, which is typically taken toward the end of the AMT school program. Also, it was assumed that students enrolled in aircraft maintenance programs will successfully obtain their A&P/AMT certificate with the intent to pursue a career in aircraft maintenance and therefore are representative of the future aircraft mechanic population.

Not all limitations were anticipated, but some were encountered as discussed in more detail in Chapter V. In the event of the success criteria not being met due to these limitations, it was determined that changing the nature of the study from confirmatory to exploratory by applying the Exploratory Factor Analysis (EFA) statistical method instead, would be an adequate alternative approach. EFA allows interpretation of the data by looking for new relationships between variables measured in order to make meaningful conclusions.

Generalizability of EFA results based on the adequacy of sample size are addressed through statistical tests, and in the case of this study, showed to be adequate. Generalizability based on sample diversity and participant demographics is assessed by making comparisons with the target population characteristics. The generalizability of this study's results was limited based on the fact that about 46% percent of final study analysis data came from ERAU students. The final study participants were not required to provide their school information; and the population data for comparing participant age and VR experience level was non-existent. Steps taken to mitigate or address this issue for the purposes of showing improved generalizability included considerations for the volunteered data by participants that represented multiple AMT schools nationwide, and a general comparison of the pilot study participant demographics with the final study.

The construct validity of the study was assumed to be adequate based on the use of a theoretical framework (i.e., TAM) for measuring technology acceptance. In other words, it was assumed that the operational variables measured through structured survey questions actually represented the conceptual variable – behavioral intention to use VR technology for maintenance training. While actual use was not measured as a variable, it could be inferred based on the measured variables, as suggested by TAM, if data on the actual use of VR for aircraft maintenance training becomes available in the future. TAM is not context dependent, and therefore, it is limited in explaining the reasons for participant responses (e.g., why participants find the technology to be easy to use, useful, enjoyable, etc.).

Finally, it was assumed that the selected statistical method ensured statistical validity and that the results are reliable using the survey methodology, with respect to test-retest accuracy. The success criterion (a sample size of 444) for response rate to the survey was assumed to be achievable and adequate for the desired power level of 0.8 (see Chapter III for additional details).

Summary

Chapter I introduced the topic of VR for aircraft maintenance and how it could be more effective for delivering training. The chapter also addressed the apparent gap in literature with respect to the lack of studies on the acceptance of VR technology by the primary users (student aircraft mechanics), and how this study could contribute to the wide body of literature as well as have practical significance within the aviation industry.

The theoretical framework of choice, the extended Technology Acceptance Model (TAM), was discussed within the scope of establishing the hypothesis and predicting actual system use. The model suggests that student mechanics' perception of VR, in terms of its ease of use and usefulness, influences students' attitude toward the technology, and in turn, their behavioral intention of using the technology for aircraft maintenance training. The chapter also covered study delimitations, limitations, and assumptions.

Chapter II and Chapter III focus on the literature review and study methodology, respectively. Chapter II focuses on the knowledge gap that the study addresses, provides an overview on VR and its many uses, discusses the development and validation of the Technology Acceptance Model, and details the findings from related studies. The chapter is intended to support the need for the study, provide justification for the theoretical model (extended TAM), and explain the basis of the hypothesis and the selected external variables. Chapter III provides the details of the methodology, explaining the study design, procedures, and the statistical approach for data analysis. The chapter also describes how the reliability and validity of the study were addressed and the appropriateness of the selected statistical analysis method, Structural Equation Modeling (SEM). The chapter is intended to provide a sufficient level of detail for the purposes of replicability and for demonstrating high rigor.

Definition of Terms

Augmented Reality	"An overlay of computer-generated content on the real world that can superficially interact with the environment in real-time" (Irvine, 2017).

Behavioral Intention	The degree to which an aircraft mechanic student intends to adopt VR technology for aircraft maintenance training (Huang & Liaw, 2018).
Extended Reality	A term applied to "all real-and-virtual virtual environments generated by computer technology and wearables" (Irvine, 2017).
Mixed Reality	"An overlay of synthetic content that is anchored to and interacts with objects in the real world – in real time" (Irvine, 2017).
Perceived Behavioral Control	The extent to which an aviation student feels able to control using VR technology for aircraft maintenance training (Fussell, 2020).
Perceived Ease of Use	The degree to which an aircraft mechanic student believes that using VR technology for aircraft maintenance training would be effortless (Huang & Liaw, 2018; Park, 2009).
Perceived Enjoyment	"The degree to which using VR for aircraft maintenance training "is perceived to be enjoyable in its own right apart from any performance consequences that may be anticipated" (Fussell, 2020).
Performance Expectancy	The degree to which a student believes that using VR for aircraft maintenance training will improve

	his/her performance as compared to traditional learning methods (Fussell, 2020; Lewis et al., 2013).
Perceived Health Risk	The degree to which an aircraft mechanic student perceives the use of VR technology for maintenance training to be posing a health risk (Tang et al., 2019).
Perceived Usefulness	The degree to which an aircraft mechanic student believes that using VR technology would be beneficial to his/her training (Huang & Liaw, 2018; Park, 2009).
Self-Efficacy	The degree to which an aircraft mechanic student has confidence that he/she is able to operate VR technology for aircraft maintenance training (Huang & Liaw, 2018; Levy & Green, 2009).
Virtual Reality	"A fully immersive, 3-dimensional, digital environment" (Fussell, 2020). "These could be created using purely real-world content (360 Video), purely synthetic content (Computer Generated), or a hybrid of both (Irvine, 2017).

List of Acronyms

3D	3-Dimensional
A&P	Airframe & Powerplant

AGFI	Adjusted Goodness of Fit Index
AMOS	Analysis of Moment Structures
AMT	Aircraft Maintenance Technician
AR	Augmented Reality
AVE	Average Variance Extracted
BI	Behavioral Intention
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CFR	Code of Federal Regulations
CG	Computer Generated
CMIN	Chi-square value in AMOS
CR	Construct Reliability
df	Degrees of Freedom
IATA	International Air Transportation Association
FAA	Federal Aviation Administration
GFI	Goodness of Fit Index
HMD	Head-Mounted Display
MR	Mixed Reality
NFI	Normed Fit Index
PBC	Perceived Behavioral Control
PHR	Perceived Health Risk
PE	Perceived Enjoyment
PEXP	Performance Expectancy

PEU	Perceived Ease of Use
PU	Perceived Usefulness
RMSEA	Root Mean Square Error of Approximation
SE	Self-Efficacy
SEM	Structural Equation Modeling
SPSS	Statistical Package for the Social Sciences
TAM	Technology Acceptance Model
TRA	Theory of Reasoned Action
VR	Virtual Reality
XR	Extended Reality

Chapter II: Review of the Relevant Literature

Chapter II presents a review and synthesis of relevant literature in support of the study, the selected theoretical framework, and hypotheses. First, virtual reality (VR) technology and its uses and benefits are discussed with specific focus on training and education within various contexts. A historical perspective of VR technology is then presented in terms of how the technology has evolved over time and how it has been applied in other industries. Next, Technology Acceptance Model (TAM) is presented along with a justification for selecting it as the ground theory of the study. Finally, relevant research and major findings are discussed, presenting gaps in research and support for the hypotheses with emphasis on VR benefits in higher education and training environments.

Virtual Reality Applications and Benefits

Background

VR technology has undergone significant improvements in recent years, enabling users to have a fully immersed experience and interactions within a virtual environment that feels highly realistic (Huang & Liaw, 2018). Perception of reality is through human senses. VR works with human senses by replacing real-world visual and sound cues with computer generated/synthetic cues, which the brain interprets as real (Jerald, 2016). As such, VR has the potential to be a very effective and powerful tool, if implemented with human-computer interaction principles in mind, based on the context of use.

The first study recorded with a head-mounted display (HMD) depicting synthetic images dates back to the 1960s, but the maturation of VR technology to what it can achieve today happened very slowly, mainly due to the limitations of computing power to process 3D images in real time (Grabowski, 2021). It took decades for VR technology to evolve and become marketable and widely used, first within the entertainment/gaming industry, and eventually, in scientific research (Grabowski, 2021). Even though the concept of viewing 3D images dates back to the 1800s with the invention of the stereoscope and simulated environments with the use of technology were achieved by the 1960s, the term *virtual reality* did not exist until the 1980s. The term was first coined by Jaron Lanier, the founder of VPL Research, who developed VR technology gear (HMD and gloves) to enable a "virtual reality" experience, as he called it (Franklin Institute, 2021). Other major milestones toward the development of VR included: a) The Sensorama (in the 1950s), a simulation of a city environment experienced by "riding" on a motorcycle and sensing the environment through 3D visual, auditory, and tactile cues (Franklin Institute, 2021); b) The Telesphere Mask (in the 1960s), an HMD for stereoscopic 3D viewing (Virtual Reality Society [VRS], 2017); and c) The Ultimate Display (in 1965), the first HMD that provided a virtual experience through the use of computer hardware and the ability of users to interact with objects in the virtual environment (VRS, 2017). The first commercialized VR gaming technology was introduced in 1991 by Virtuality Group who produced arcade games/machines. Since then, the gaming industry has invested billions into maturing VR (Clavin, 2019), which has enabled scientists, educators, and manufacturers to leverage the technology and apply it as a tool for data analysis, training, engineering, and production; however, the pioneer work and interest in VR by scientists, engineers, and government sectors, including the military, dates back to as early as the 1960s with the development of Furness' Flight Simulator (VRS, 2017). Furness' work continued through the 1980s and resulted in the development of the Super Cockpit, intended for pilot training. The Super Cockpit enabled pilots to interact within a simulated environment in real time and provided a way of tracking and integrating movement and aircraft control (VRS, 2017). Around the same time, the National Aeronautics and Space Administration (NASA) developed Project VIEW (Virtual Interface Environment Workstation), a VR technology intended for astronaut training. The technology was distinguished from others by featuring gloves for advanced haptic interaction (VRS, 2017).

With optical and computer technology advancing fast, VR started becoming popular and available within commercial markets in the 1990s. Other uses of the technology were also identified, such as for the treatment of Post-Traumatic Stress Disorder in veterans and for map services, such as 3D street viewing; however, the main surge in VR products happened in 2016 and later (VRS, 2017). Within the last five years, VR has become increasingly accessible and relatively affordable, resulting in more use than ever in government and private sectors, beyond just fun and games. One such example is the joint effort by the Center for Data-Driven Discovery (CD3) (a non-government VR research group), Caltech (a privately supported science and engineering institute), and NASA (a government agency) (Clavin, 2021). The three entities have partnered to develop and implement multiple VR tools for scientific investigation and learning/training purposes. In one case, VR is used to detect tumors that could go undetected with current medical screening methods. The technology is being combined with machine learning/artificial intelligence to help medical professionals detect potential malignancies more easily. This application of VR technology is one of many that CD3 and Caltech are working on for future implementation, which could potentially transform the scientific paradigm of how investigations and data analyses are performed (Clavin, 2021).

VR Applied to Training and Education

Within the last decade, VR applications have been used by government and commercial sectors to enhance training in a variety of areas, including military, medical, flight training, and education. Additionally, the technology has been applied within engineering, space, and manufacturing industries (Vaughan et al., 2016). VR is deemed especially useful for physical tasks or procedural training because it can create an illusion of spatial presence (Grabowski, 2021). It also supports building and/or exercising muscle memory through real movements in a virtual environment and has the potential to improve decision-making by presenting the trainee with consequences of their actions within the simulated environment (Grabowski, 2021). Based on the constructivist view of instructional theories, using VR for education and training can lead to higher motivation to learn because the technology provides an opportunity to engage in real-life activities (Huang & Liaw, 2018). In its advanced form, VR can be self-adapting to suit the individual needs of each learner, such that training/education scenarios are constantly tailored based on the learner's responses. This is achieved through feedback and control loops of adaptive systems, which enables the machine learning phases of prediction, recognition, detection, and optimization (Vaughan et al., 2016). These known and promising capabilities have generated a number of research studies on the training effectiveness of VR.

Current uses of VR technology in military settings include development and implementation of training, both for flight and ground missions. The Air Force Agency for Modeling Simulation presents a successful example of such efforts because VR is being used to change the landscape for air combat training to minimize cost and eliminate risks associated with real air combat training (Lange, 2020). Traditional air-to-air combat training requires large ranges to practice in, and it is highly risky and challenging to simulate the type of threats that are anticipated from enemies now or in the future during warfighting (Lange, 2020). Hence, VR provides a safer alternative for simulating those threats within a synthetic environment and training military personnel to respond to threats within a controlled physical environment; in real life, this environment entails having to make risky maneuvers and pushing the limits of the aircraft (Lange, 2020). Therefore, VR training is currently being promoted and advanced by the Department of Defense for more effective training as well as for the purposes of addressing the limitations of the traditional training methods (Lange, 2020) and in turn avoiding hazards and potential accidents during training. Similarly, VR is currently used for training ground military crews, during which the technology provides an opportunity for implementing scenarios that incorporate lessons learned from real battles and transferring knowledge to younger soldiers. Compared to traditional military training, VR training enables the military to study the individual behaviors of the trainees, how they respond to emerging threats, and how they make decisions. The ability for the military to extract this data more easily and accurately from VR training performance ensures that the best trained soldiers are selected for the real mission (Kozlak et al., 2013). While the current use of VR in military training provides many benefits, such as allowing soldiers to make and learn from mistakes, allowing them to repeat training as many times as needed, providing a safer environment, reducing the cost of training, and enabling behavior analysis, it does have some drawbacks to consider. Typical VR training simulators used by the military today are in a temperature-controlled environment. The trainee soldiers engage in a realistic combat task, but without other environmental factors that could influence their performance, such as heat and cold, the weight of gear that they would be carrying in a real combat situation, and so on. Also, from a psychological aspect, being able to go through walls in VR without collision issues may lead to perceiving the training as less serious (Kozlak et al., 2013).

VR technology has also been shown to provide effective training in the law enforcement field. VIRTSIM by Raytheon (in partnership with Motion Reality Inc.) is one such commercially available product, tailored toward public safety. The primary users are law enforcement agents, border security officials, U.S. Marshals, FBI agents, and other first responders. The training is designed to be as realistic as it can be with participants carrying realistic weapons and not limited in their tactical movements within the physical training environment (Raytheon, 2012). VIRTSIM training produces the added advantage of physiological responses that would be expected during actual missions, such as sweating and increased heart rate (Raytheon, 2012), because of the intense immersive experience that VR technology provides.

Outside of defense, VR/AR technologies are already standardized as tools used for training, production, design and development, and other operations. The technologies have comparable benefits associated with risk/loss reduction, enhanced learning/knowledge capture/skill development, efficiency, and reliability/quality improvements (Kozlak et al., 2013). Both technologies are currently used by many companies that deal with environmental hazards as part of their daily operations and/or require skill-based training. VR/AR enable users to conduct a walkthrough of hazardous facilities without being exposed to the actual hazards and effectively learn the skills required to operate and maintain a complex system. For example, the Exxon Mobile company uses simulation training for the operation of vessel compressors with over 100 pressure valves that have various functions. The intent of the training is not only to learn how to operate the system but also how to maintain it or manage failures of the system (Kozlak et al., 2013).

VR Benefits and Implications on Aircraft Maintenance Training

The benefits of training with virtual reality (VR) technology over traditional training methods are well established in the existing literature. VR enables a whole-body activity and higher levels of engagement which leads to learning gains (Lindgren et al., 2016). This is especially evident for physical and skill-based tasks, such as aircraft maintenance, flight training, design for maintainability, assemblability, manufacturing, and production. Many studies have demonstrated VR training benefits in terms of error prevention, effectiveness, and efficiency. A discussion of research studies and other literature focusing on VR benefits within the context of training and education is provided below, along with implications for aircraft maintenance training, specifically.

A study conducted by Jang et al. (2017) employed a case study methodology to investigate VR training for medical procedures. The participants were medical students, tasked with studying complex anatomical structures by using VR with the ability to directly manipulate objects in a simulated environment. The study concluded that VR enhances learning beyond passive viewing. Also, the authors concluded that VR is especially useful for students with low spatial abilities because it puts the student in a fully immersed virtual environment that is very realistic to the actual life situation. Gains associated with direct manipulation in VR have also been demonstrated by Marzano et al. (2015) for an aircraft carriage maintenance task. The ability of participants to manipulate every single part of a complex assembly was shown to be value added. Traditional textbook/video methods do not provide this opportunity, whereas field training does, but the cost and potential risks associated with hands-on training is high for complex aircraft maintenance tasks.

A study conducted by Makowski et al. (2017) focused on the cognitive aspect of simulation learning. Their study explored the question, Does *being there* and reacting to a stimulus as if it were real enhance memory encoding? The results were interesting in that higher emotional experience and sense of presence in a simulated environment was found to enhance factual memory but not temporal order memory. In other words, the sense of presence helped participants focus their attention on the stimulus and lead to better memory encoding, which would help with learning a given task, but not with learning procedural/sequential steps (Makowski et al., 2017). The study has some implications on the use of VR for aircraft maintenance training because it requires consideration of the objective of the training. If the objective is to teach student mechanics a procedural and sequential step, then VR training may not be suitable to help them remember those steps; however, if the objective is to teach students how a specific task is achieved, then VR training would bring significant benefits.

Finally, many other studies have investigated how VR training improves performance in terms of task efficiency and error reduction. In Burigat and Chittaro's (2016) study, participants were tasked to navigate within a given environment.

Participants either used a virtual environment (VE) or physical maps/diagrams to navigate. Those who used a VE showed a better spatial awareness for their relative position within the environment and performed better in a subsequent virtual evacuation task. Participants also rated the VE approach to be more enjoyable and easier to comprehend (Burigat & Chittaro, 2016).

Langley et al. (2016) were able to show that those who received VR training in automotive assembly operations made fewer errors when actually performing the task, as compared to those who received traditional training. These results are consistent with the findings of Cooper et al. (2016) in which VR training with augmented cues led to better physical and skill-based task performance (lower task time and lower errors compared to those who did not receive VR training ahead of completing a tire change task). The results suggest that one of the benefits of using VR for training is to provide familiarity. Similar results were observed in a study by Brauer and Klingauf (2008) with 100 Lufthansa pilot participants. Participants received VR flight training with a virtual captain, which reduced task familiarization time and facilitated learning.

In summary, performance of aircraft mechanics can be influenced by their environment, making the task at hand more/less challenging from an accessibility and body posture perspective. Spatial knowledge gained through VR becomes an important factor in teaching aircraft maintenance students, in that it can provide a realistic experience and potentially lead to better performance in the field. While the research on VR benefits is scattered across various industries, the benefits identified in these studies are transferable to similar tasks within the aviation industry, to include aircraft maintenance, because aircraft maintenance also requires manipulation of components within a given space and procedural knowledge. Combined with the successful examples of how VR is already used within other industries as a standard method of training, full deployment of VR technology for aircraft maintenance training and certification is only a matter of time. Hence, it is important to understand how aircraft mechanic students perceive VR today to make sure that it is implemented properly in the future to satisfy user needs and to change negative perceptions, if any exist. This study addresses a critical issue pertaining to the adoption of VR technology in the aircraft maintenance training by investigating the acceptance of the technology by the students, capturing their behavioral intentions/attitudes with practical near-term implications.

Current State of VR & Technology Adoption

VR technology is continuously and rapidly advancing to bring more benefits and dynamic training options. Years of research on effectiveness and efficiencies of VR has resulted in extensive investments by the aviation industry to adopt the technology. It is projected that the investment will surge by 2030 to over $17 billion, compared to the $1.76 billion in 2023 (HQSoftware, 2024). The aviation industry struggles with the shortage of aircraft maintenance technicians, which is attributed to the high cost of AMT programs and limited opportunities the traditional training methods present to students. AMT schools are experiencing a 30% drop in graduates on a yearly basis. The aviation industry views VR as a way of addressing some of these challenges by providing students with a more effective, efficient, inspiring, and seamless training opportunities (HQSoftware, 2024) (CAE, 2024).

The benefits of VR specifically for aircraft maintenance training have been demonstrated through many studies, by the U.S. military, and AMT schools. According to a CAE (2024) report, a study on aircraft maintenance training using VR showed that students completed training four times faster with VR compared to the same training provided in a traditional classroom setting, they were four times more emotionally engaged with VR, and 275% more confident in their skills after VR training (CAE, 2024). Hill Air Force Base (AFB) has seen similar benefits with their implementation of VR for aircraft maintenance training of new recruits and working veterans (Cromar, 2024). Trainees have shown improved training retention of more than 35% with VR, compared to 10% with traditional methods. With VR providing effective familiarization of different aircraft types/sub-systems/parts, trainees were able to perform maintenance tasks successfully and efficiently on aircraft that they have never seen before in real life (Cromar, 2024). In 2022, the Advanced Manufacturing Center was the first Part 147 AMT program in the U.S. to pursue approval to use VR for satisfying FAA training requirements for aircraft painting and coating. Their VR technology is the same one used at Hill AFB. VR enabled this AMT school to accurately measure student performance based on material use and cost, thickness of paint applied, the techniques students were using to apply the paint, any overspray issues, etc. (Marshall University, 2021).

The current state of VR is already contributing significantly to the aircraft maintenance training needs within the aviation industry; and integration with other technologies in the future, such as Artificial Intelligence (AI) would only improve the training by allowing it to be more tailorable and personalized with targeted practice and remedies for building specific knowledge and skills (HQSoftware, 2024). Still, there are many challenges to consider for aircraft maintenance training with VR. Some of these challenges include resolution, latency, the complexity of the technology requiring highly specialized professionals to maintain and troubleshoot, the effort required for content development, and technology acceptance by students (HQSoftware, 2024) (Murugan, 2024). According to Li (2023), the simulated environment, learning content, prompts and error correction, and task-based evaluations are the design parts that need to be considered during VR technology development for aircraft maintenance training. Currently, there is no standardized approach to each of these design parts, which means that implementation can vary significantly from one AMT institution to another. While the FAA acknowledges VR benefits and is working towards creating standards for compliance with training requirements using VR, with the aviation industry traditionally erring on the more conservative side, it will take some time and rigorous testing of VR to achieve widespread adoption (Murugan, 2024). More recent research studies on aircraft maintenance training focus on VR implementation, and how it could be optimized to enhance learning and influence the aircraft maintenance students' adoption of the technology.

With respect to VR technology acceptance for aviation training, enjoyment, usability, and usefulness are common themes in many studies; and establishments that already use VR for training are making continuous efforts to measure those factors and improve the curriculum as necessary. For example, the Florida Technical College of Aeronautics is using VR for their Aircraft Systems course, focusing on providing an overview of the Boeing 737 NG and various systems of other commercial airplanes. This school's VR is integrated with an eye tracking technology as students interact with flight deck displays and controls to complete procedures (Lowenstein, 2024). Continuous student and instructor feedback was an important part of this school's journey to modernize their classroom with VR. With positive feedback collected during the initial demonstrations of VR in 2023, the school proceeded with surveying their students in 2024 on VR usability, enjoyment, and usefulness for learning aircraft systems that are advanced (Lowenstein, 2024). A study by Gomez-Cambronero et al. (2023) attempted to address the challenge of VR acceptance with a *gamification* approach. The idea was to incorporate the same elements that make games enjoyable into a VR scenario in order to make the aircraft maintenance training more appealing to students. The results showed that gamification of aircraft maintenance training with VR motivated students to complete more practice exercises, and that students found this type of implementation to be easy to use (Gomez-Cambronero et al., 2023).

Recent research efforts by academia and AMT programs are on the right path for optimizing VR implementation to promote independent and enhanced learning and to achieve technology acceptance. However, these efforts are very focused on a specific implementation, produce results that are often maintenance scenario or task-specific and apply to a limited population (i.e., own students). It is not always the case where research is able to demonstrate aircraft maintenance students' acceptance of VR technology. For example, a study by Starr (2024) considered a typical academic framework for aircraft maintenance training, which includes learning through acquisition, inquiry, discussion, and practice. Researchers created and tested a virtual classroom prototype using VR to support the natural progression that students need to make through those phases and types of learning, with an opportunity for an instructor to enter the virtual environment to provide additional guidance to students. With this implementation of VR for aircraft maintenance training, about 65% percent of participants expressed a preference for the traditional training methods (Starr et al., 2024). Most recent research strives to establish VR technology acceptance for aircraft maintenance training with specific implementation and with specific user groups, substantiating the selected factors for this study. However, there is still much to learn about the attitudes towards VR by the aircraft mechanic student population as a whole.

Theoretical Framework of the Study

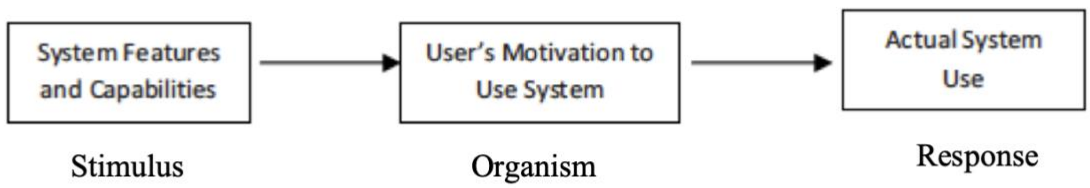
Previous sections included an overview of VR technology use and benefits in training and educational environments. The review of relevant literature provided substantial support for studying and implementing VR for aircraft maintenance training, as the technology has shown to be effective in similar applications, with many of the learning benefits discussed being transferable. However, one important factor that the previous sections did not cover is the adoption of VR technology for training, and the factors that influence the students' behavioral intention to use the technology or not. Studying the behavioral intention of using VR within an education or training context requires a theoretical framework and validated methodology in order to identify and examine factors that influence such decisions; therefore, this section includes a detailed discussion of TAM, the theoretical model of choice for the proposed study.

Technology Acceptance Model

The Technology Acceptance Model (TAM) provides a validated theoretical framework for hypothesis testing and is widely used by researchers to study technology adoption. The model was developed by Fred Davis at the Massachusetts Institute of Technology (MIT) to address the rising need in research during the 1970s for predicting technology acceptance of information systems. The basic concept behind TAM suggests that the actual behavior in accepting/rejecting a technology is a motivation-driven response, impacted by system features and capabilities, as shown in Figure 2 (Chuttur, 2009).

Figure 2

The Relationship Between Stimulus-Organism-Response

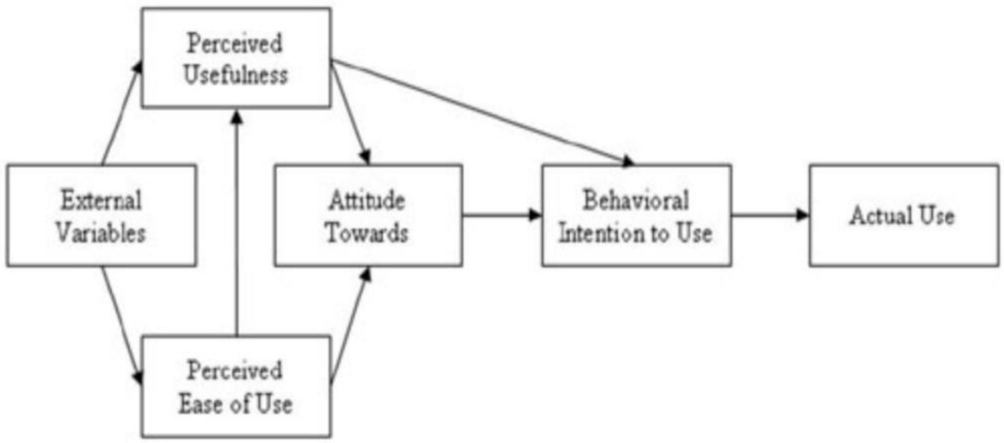


Note. The relationship between Stimulus – Organism - Response as an explanation for technology acceptance. Adapted from "Overview of the Technology Acceptance Model: Origins, Developments and Future Directions," by M.Y. Chuttur , 2009, *All Sprouts: Working Papers on Information Systems, 9*(37) (https://aisel.aisnet.org/sprouts_all/290).

The TAM is an extension of the Theory of Reasoned Action (TRA) model, which introduces the variables of attitudes, subjective norms, and intentions as predictors of technology use. In comparison, TAM adds two additional variables, Perceived Usefulness (PU) and Perceived Ease of Use (PEU), as directly influencing attitude. It also accounts for external factors that may have indirect influences (Mishra et al., 2014). Figure 3 shows a modified early version of TAM.

Figure 3

Early Modified Version of TAM

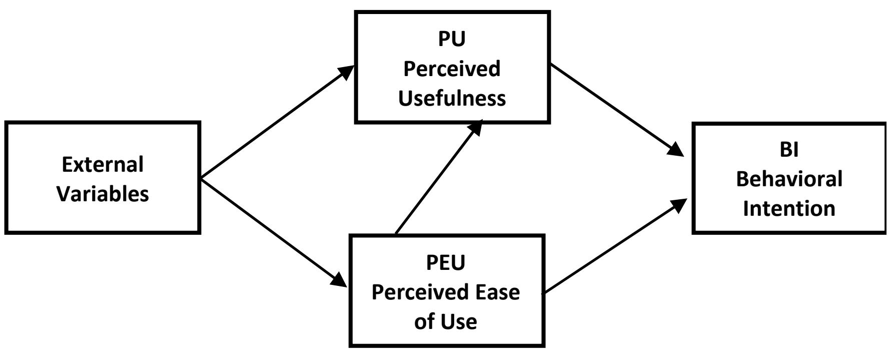


Note. Early modified version of TAM showing the influence of perceived usefulness and perceived ease of use on attitude and/or behavioral intention. Adapted from "An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use E-Learning," by S.Y. Park, 2009. *Educational Technology & Society, 12*(3), 150–162.

As part of early validation efforts, Davis defined, operationalized, and tested the two variables of PU and PEU and the originally theorized relationships (Chuttur, 2009). PU was referred to as *the degree to which an individual believes that using a particular system would enhance his or her job performance* and PEU as *the degree to which an individual believes that using a particular system would be free of physical and mental effort* (Chuttur, 2009). The variables were operationalized and tested through ten-item psychometric scale measures. The reliability and validity of TAM was demonstrated in a study with 112 employee participants from IBM, where the actual use of technologies studied was found to be strongly correlated with PU and PEU. After confirming the statistical relationships and based on additional findings since the development of TAM, Davis revised the model in 1993 to depict the direct relationship between PU and actual use of technology, and the direct relationship between technology characteristics on user attitudes (Chuttur, 2009). After further studies, Davis introduced the construct variable of Behavioral Intention, which was theorized to be indirectly influenced by PU and PEU. However, based on additional findings, Davis later removed the variable of attitude toward technology, directly mapping PU and PEU to Behavioral Intention (Chuttur, 2009). This became known as the final 1996 version of TAM (see Figure 4).

Figure 4

Final 1996, TAM Model as Applied in this Study



Note. Adapted from "Overview of the Technology Acceptance Model: Origins, Developments and Future Directions," by M.Y. Chuttur , 2009, *All Sprouts: Working Papers on Information Systems*, *9*(37) (https://aisel.aisnet.org/sprouts_all/290).

The final 1996 version of TAM was studied in terms of its limitations and application to different settings. TAM was found to be consistent in predicting technology acceptance with high test-retest reliability. The methodology used in the studies to investigate the reliability and validity of TAM was criticized by some researchers based on the fact that the data was subjective, the results were not generalizable due to participants being students, and the technology the model was tested on was not for mandatory use but rather voluntary. Also, it has been suggested that the model is missing the affective/cognitive attitudes – potentially meaningful variables when it comes to technology adoption (Chuttur, 2009).

Currently, there are many technology acceptance models, but some of them like the TRA are context specific. In other words, user beliefs must be evaluated within the context of a given task, specific environment, and implementation of the technology. In a 1989 study, TAM was compared to the TRA model and found to be easier to apply and less costly because it is a technology independent model. TAM was also compared to the Theory of Planned Behavior (TPB) model, which is similar to TRA, but includes additional variables (e.g., perceived behavioral control). While TRA was found to be just as effective in predicting technology use, the model is more complex (Chuttur, 2009). Therefore, over the years, TAM has become one of the most popular models as a simple and suitable model for studying any technology in any environment. For this reason, the final 1996 version of TAM, extended with the external variables, was the model of choice for this study.

Relevant Research and Major Findings

This section focuses on the most relevant literature and major findings for the purposes of establishing the need for this study and in support of the research questions/hypotheses. The search for relevant literature was conducted by using various combinations of key terms, such as *aircraft maintenance*, *virtual reality,* and *technology acceptance model*, etc., using the databases accessible through ERAU's Hunt Library and the Google search engine. The search was limited to peer reviewed and scholarly articles published within the last eleven years (2010 to 2021). With each search result (sorted based on relevance), the first 30 articles were reviewed to identify studies related to VR benefits and adoption for training. Additionally, one literature source was used from a recent dissertation published in the ERAU Scholarly Commons. The literature search produced a total of 11 articles that related to the scope of the study, all of which were considered in establishing a need for the study, as support for the selected methodology, and to develop the hypotheses. The review of relevant research showed that many studies substantiate the claim of VR technology providing gains in learning and training, but literature specific to VR use in aviation is limited and scattered across various disciplines, such as design for maintainability/assemblability and pilot training. Even more limited is research on the use of VR for aircraft maintenance training. Of the 11 relevant studies identified, only one, conducted by Sagnier et al. (2020), investigated VR use and acceptance for an aeronautical assembly task. As for literature on acceptability of VR for aircraft maintenance training, using the Technology Acceptance Model (TAM) as the theoretical framework, it is non-existent based on the literature review conducted.

Of 11 studies identified as relevant, four have strong implications for this research based on task or concept similarity. The four studies demonstrate a similar methodological approach followed by researchers to predict VR or AR technology acceptance in various applications and settings, reveal relationships between the variables, and establish the basis of the hypotheses in support of this study. This section provides a discussion of these four studies as the most relevant literature and the significance of the researchers’ findings.

The Technology Acceptance Model (TAM) was applied by Sagnier et al. (2020) to study VR for an aeronautical assembly task, which included a set of short aircraft manufacturing operations to rivet components. The participants were 89 engineering students randomly assigned to two devices, CAVE vs. Head Mounted Display (HMD) VR. TAM was extended to include User Experience (UE) variables to investigate relationships between pragmatic and hedonic quality-simulation (cybersickness, presence, and individual/personal innovativeness). The researchers were interested to see the relationship between these factors and Perceived Ease of Use (PEU), Perceived Usefulness (PU), and Intention to Use (ITU). The study revealed that VR use is influenced by PU (positively) and by cybersickness (negatively) (Sagnier et al., 2020). These findings support the selection of external variables and the hypotheses – the positive relationship between PU and Behavioral Intention (H1) and the negative relationship between Perceived Health Risk (PHR) and PEU/PU (H8, H9).

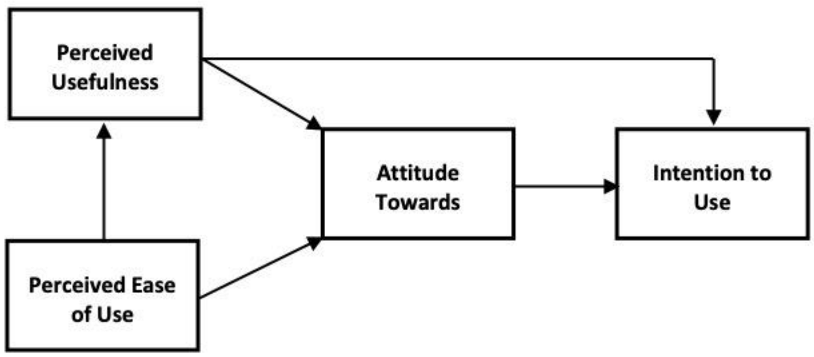
A study conducted by Junglas et al. (2013) also investigated VE/VR, focusing on the social component of information systems and how it can impact technology acceptance. The study introduced the concept of s*ociability* and Perceived Enjoyment (PE). More specifically, the results suggested that highly interactive virtual environments enabled by VR technology could lead to higher enjoyment and positively influence user perceptions and in turn acceptance of the technology (Junglas et al., 2013). The findings of this study support the selection of PE as an external variable, hypothesized to influence PU and PEU (H6, H7).

A more recent study by Fussell (2020) investigated VR technology acceptance for flight training. The theoretical model used was TAM, which was expanded to incorporate constructs from Theory of Planned Behavior (TPB) and Unified Theory of Acceptance and Use of Technology (UTAUT). More specifically and relevant to this research, Fussell (2020) hypothesized a positive relationship between performance expectancy (PEXP) and perceived usefulness (PU), perceived behavioral control (PBC) and perceived ease of use (PEU), and self-efficacy (SE) and perceived ease of use (PEU). The data was analyzed using SEM. The findings supported the hypothesized relationship between PEXP and PU and SE and PEU. While the relationship between PBC and PEU was not supported as hypothesized, based on the literature review conducted by the researcher, there was sufficient reason to include the construct of PBC in the TAM model as one of the external variables, as it has been previously associated with behavioral intent and actual use of technology (Fussell, 2020). The literature review and findings of Fussell (2020), at minimum, support the selection of PEXP, PBC, and SE as external variables that could be associated with PU and PEU of VR technology within the context of aircraft maintenance training (H4, H5, H10, H11).

Perhaps the most relevant research is by Wang et al.(2016) on AR technology acceptance for aircraft maintenance training. AR is thought to have similar benefits when used for training as it also enables a simulated interaction with enhanced cues for learning a given task. The study by Wang et al. (2016) is the only study known to date that uses TAM to understand similar-to-VR technology acceptance specifically for aircraft maintenance training. The study employed the simplified version of TAM and a survey methodology to collect data from student mechanics. The results obtained in the study support the hypothesized relationships between PU and PEU and all other core variables of TAM (H1, H2, H3). Figure 5 shows the confirmed model and hypothesized relationships from the study conducted by Wang et al. (2016).

Figure 5

Confirmed Hypotheses from Wang, Anne, and Ropp (2016)



Note. Adapted from "Applying the Technology Acceptance Model to Understand Aviation Students' Perceptions Toward Augmented Reality Maintenance Training instruction," by Y. Wang, A. Anne, and T. Ropp, 2016, *International Journal of Aviation, Aeronautics, and Aerospace, 3*(4) (https://doi.org/10.15394/ijaaa.2016.1144).

Summary

Over the past 10 years, there have been multiple studies conducted on Virtual Reality (VR) use for employee or student training, but only a few studies focused on acceptance of the technology by aircraft mechanic students. Nevertheless, these studies provide significant findings that support the scope and need of this research. There is no specific study on acceptance of VR as applied to maintenance training using TAM and for the same combination of external variables selected in this study. The most relevant studies either applied TAM to VR to the areas of manufacturing, design for maintainability, pilot training, etc., or applied TAM to a similar technology, such as AR. Many of the studies on VR technology acceptance failed to include any theoretical model, and used a limited number of participants, who often were from a single institution, which compromises generalizability of findings to the greater population.

Synthesis of relevant literature shows that another knowledge gap is in the scope of existing studies. There is no single study that examines a comprehensive set of external variables, to include perceived enjoyment, perceived health risk, and self-efficacy all in one study; and there is no consistency in the results obtained with respect to the hypothesized relationships between the variables. This makes it difficult to select/incorporate external variables into TAM and to formulate hypotheses. Finally, there is a lack of rigor in previous studies related to VR technology acceptance for training with respect to study methodology, data collection, analysis, and interpretation. As a result, some studies fail to provide a viable explanation for their findings. Much of this comes from not utilizing a clear theoretical framework, especially when researchers attempt to incorporate many different models, but end up with convoluted hypotheses and testing as observed in the study by Sagnier et al. (2020). The Sagnier, et al. (2020) study incorporated UX variables, but applied multiple different rating scales to collect and analyze the data, which made the methodology too convoluted from a statistical analysis perspective, because using different measurement scales for variables that are hypothesized to be associated with one another could potentially compromise the reliability of the results. The need for an experimental design to study acceptance was not well explained either, as the theoretical framework used by the researchers was TAM (a model that does not suggest causation but rather an extent and direction of associated relationships). Another example of a study that lacked rigor due to no theoretical framework being used to address VR technology acceptance was by Brauer and Klingauf (2008). The study was conducted with 100 Lufthansa pilot participants. The main focus was on the effectiveness of VR, but conclusions were also made on acceptance of VR for pilot training with no ground theory to justify or explain the author's interpretation of the results. For that reason, the study was excluded from the list of relevant literature and for supporting the hypotheses of this study. Therefore, this research was an opportunity to address some of these gaps with a simplified but rigorous method of applying the final 1996 version of TAM (extended), to test very few variables of interest that can be supported by previous research findings and statistically analyzed. The methodology, in terms of the number of participants, the type of statistical analysis, and use of a previously validated survey instrument, ensured trustworthy results. The study also focused on a limited set of external variables, specifically those that have been deemed to be supported through previous research. This preserves the simplicity of the theoretical model and was deemed appropriate for confirmatory analysis.

Chapter III: Methodology

Chapter III describes the research methodology for this study to establish replicability. The section addresses the research approach, design and procedures, sampling, data collection, and data treatment. The reliability and validity of the study are also discussed.

Research Approach

This study employed quantitative research methodology. Quantitative methodology is appropriate when a theoretical model is used as the bases of the hypotheses. The intent of the study was to confirm/disconfirm the hypothesized relationships through objective statistical analysis, which is best achieved and communicated through a quantitative approach. Understanding the reason behind student mechanics' perceptions was not part of the study scope and therefore, a qualitative research methodology was not selected. The theoretical framework selected, the Technology Acceptance Model (TAM), required data collection via a cross-sectional survey. A cross-sectional survey suggests that the data represents a snapshot in time. Structural Equation Modeling (SEM) was applied to the quantitative data to assess model fit and hypothesized relationships -- to make statistical inferences about the student mechanic population, more specifically their attitudes toward VR use for aircraft maintenance training. Since the study utilized TAM as the main theoretical framework, the research approach was considered deductive – confirmatory in nature. It is important to note that the study was non-experimental, in that variables were not manipulated, and no causal relationships can be established.

While TAM offers a theoretical framework that has been validated through many studies before, and the survey instrument enables the operationalization of the conceptual variables, considering the uniqueness of each study with regard to the selected external variables, the model still needs to be validated. Therefore, a pilot study was conducted prior to deploying the final study to help validate the hypothesized model and make necessary adjustments to the model and the survey instrument. The pilot study was administered to students enrolled in the Embry-Riddle Aeronautical University (ERAU) aviation maintenance science program on the Daytona Beach Campus. Since the purpose of the pilot study was to validate the survey instrument, ensuring that the questions are relevant, clear, and measure the constructs as intended, the anticipated sample size of 40-60 participants was deemed acceptable. The pilot study gave an opportunity for an iterative approach to improving the overall validity and reliability by updating the survey questions based on preliminary results. The data from the pilot study was analyzed using Confirmatory Factors Analysis (CFA) to test the relationship between variables, make conclusions on model fit, and in turn, update the hypothesized model as necessary.

Design and Procedures

The research design of this study was a survey. Survey design is a cost-effective means of collecting a large amount of data, and it is a common method used in social/behavioral sciences (Vogt et al., 2012). Surveys produce subjective, and in this case, quantitative data, which can be statistically analyzed and generalized to the population being studied. For the purposes of the study, the survey utilized a five-point Likert scale to collect and code participant responses for quantitative statistical analysis, where a rating of 1 represents *strongly disagree* and 5 represents *strongly agree*. A five-point scale was selected based on researcher preference, though it is consistent with similar studies using TAM. It is also thought to be less frustrating for participants, which in turn produces a better response rate compared to a seven or nine-point scale. Demographic information was also collected to make comparisons in student mechanic responses. The survey instrument was designed based on multiple previous studies that used TAM to investigate technology acceptance within various contexts. References for studies used to generate a preliminary set of questions for measuring the constructs are presented in Figure 6. An effort was made to adapt a survey instrument specifically from studies that applied TAM to VR (or similar technology, such as AR) and studies that examined the same set of external variables. This approach was important in developing the survey questions to ensure that the questions measured the intended construct, as demonstrated previously. The questions did require some customization to fit the context of this research. For example, if the source study applied TAM to predict VR acceptance for flight training, then the core of the question was maintained with respect to the constructs and indicator variables, but the use of VR within a given context was updated to state, "for aircraft maintenance training."

The study applied Structural Equation Modeling (SEM), a multivariate statistical analysis technique, to analyze the quantitative data collected from the survey and to make inferences about the target population. SEM is considered appropriate for studies that utilize a theoretical model, such as TAM, and are confirmatory in nature. It has been used in many previous studies to analyze relationships between indicator variables and latent variables as proposed by a given theoretical model. In the study and per TAM, the indicator variables were perceived usefulness (PU) and perceived ease of use (PEU), and the latent variable was the behavioral intention (BI) of using VR for maintenance training. However, as a parametric statistical analysis technique, SEM does require interval-quality data, which in this case was achieved through the summation of the Likert scale responses across multiple items/questions that measure the same variable (Fussell, 2020).

SEM involves a two-stage analysis: (1) Confirmatory Factor Analysis (CFA) and (2) Structural Equation Modeling (SEM). The model utilizes regression equations to analyze relationships that exist between the variables from sets of data, if any (Byrne, 2016). CFA is a measurement model, intended to investigate relationships between construct variables and their observed variables. In other words, it is conducted to confirm the theoretical model, in this case the extended TAM used in the study. CFA method looks at covariances of latent variables and their associated measured variables to confirm or disconfirm theorized links, which ensures the validity and reliability of the measures and allows testing for SEM assumptions. The process of conducting CFA requires data to be prepared (e.g., replaced, imputed, transformed as necessary) (Truong, 2018). CFA is the first statistical analysis applied to the data to establish valid relationships and good model-fit, followed by SEM to test the significance of those relationships.

The second stage analysis, SEM, is intended to test the entire structure of the theoretical model, the direction of the relationships hypothesized, and the strength of those relationships. The analysis produces results that are interpreted similar to CFA (Truong, 2018). The process of conducting SEM includes creating a path diagram based on the confirmed theoretical model that is being tested. Factor loadings, model-fit indices, and MI values for post-hoc analysis were calculated. The MI values could reveal new relationships that were not initially hypothesized, in which case an additional literature review would be required to explain those relationships. The main goal of SEM is to make conclusions on model-fit, which means assessing if the data substantiates the proposed model. It is reasonable to expect a good model-fit in SEM (stage 2) analysis, if the CFA has revealed a good model-fit to begin with (Truong, 2018).

In summary, the study steps included (a) defining the construct variables, (b) developing a survey instrument, (c) conducting a small-scale study (pilot study), (d) revising the model and survey instrument as suggested by the small-scale study results, (e) conducting a large-scale study, and (f) analyzing the data through full SEM (assuming required sample size is achieved). The definitions of construct variables and the survey instrument were derived from studies that utilized TAM or studies that utilized other theoretical frameworks to define and investigate the same variable as in this study. Hence, the survey questions were also based on previous literature for higher validity, though some customization/tailoring was required to fit the context of the study. The small-scale study was used refine the survey instrument and hypotheses. The results of the small-scale study helped determine if adjustments to the model with respect to the construct variables were needed. The finalized survey was administered digitally within a set timeframe, and data collected and analyzed in two stages, as described previously. More procedural details are provided in the following sub-sections.

Apparatus and Materials

There were no physical materials used in the study, other than the incentive being offered as part of a raffle and a computing device to digitally administer the survey and to collect and analyze the data. The survey instrument was developed, and its link distributed to participant emails using a web-based application, Google Forms. The survey included a briefing and introduction section that explained the study's purpose, instructions, and participant rights. To ensure that all participants had a basic understanding of VR, a short Power Point presentation with a video was shown before the survey began. The initial set of questions obtained demographic data, followed by questions that were intended to measure the variables in the study. The data from the survey was analyzed using the SPSS AMOS software, SPSS, and Microsoft Excel.

Population/Sample

The target population for the study was students currently enrolled in aircraft maintenance schools within the U.S. The sampling frame was primarily driven by the list of FAA approved maintenance schools but was also limited by the institution's willingness to support the study, allowing survey administration to their students and access to/availability of student email addresses. When this study was planned, there were 177 FAA approved schools across the U.S. (FAA, n.d.-a). For practical reasons, judgement sampling as a form of convenience sampling, a non-probabilistic strategy, was used to target institutions that were more willing to support the study or institutions with higher and lower enrollment to achieve some level of diversity. Both public and private schools were considered. All 177 FAA approved schools were contacted via phone/email/social media and extended an invitation to participate.

The size of the target population is the total number of students enrolled across all FAA approved maintenance schools. Since enrollment is subject to change at any given moment, and the FAA does not require schools to report current enrollment data, it is difficult to calculate the population size. Based on a business aviation article published by Broderick (2017), the population size was 18,000 students as of mid-November 2017. Anticipating an increase in enrollment since then, the population size was estimated to be 20,000 to 50,000 at the time of this research. Determining the sample size for SEM is a challenge acknowledged by many researchers (Wolf et al., 2015). Compared to other statistical techniques, SEM is more sensitive to sample size and requires a larger sample to ensure accurate and valid conclusions about a given population. The sample size required for SEM analysis also depends on the number of variables being studied. A theoretical model with more variables and complex relationships would require a larger sample size (Fussell, 2020). Other factors to consider when calculating the sample size include the desired confidence level (95% vs. 99%) and the margin of error (typically set at 5% maximum). When the population data is available, the sample size can be determined mathematically, which is the most preferred method. One such mathematical method is using the Raosoft (2004) sample size equation, as shown below:

$$x = Z(c/100)^2 * r(100-r)$$

$$n = Nx/((N-1)E^2 + x)$$

$$E = \mathrm{Sqrt}[(N-n)x/n(N-1)]$$

In this case, the exact population size data was not available. Attempts were made to collect the enrollment information by contacting the schools directly, which was unsuccessful due to the schools' failure/inability to provide the information. Hence, the planned workaround technique of estimating the required sample size based on similarity of schools, to allow mathematical calculation of sample size and to validate the success criteria of the study, was also not possible.

Alternative approaches for defining the sample size were considered. For example, using the guidelines provided by Hair et al. (2010) and assuming model complexity equivalent to seven or more constructs, the minimum sample size would be 500. This is consistent with the mathematical estimation of 375-655 survey responses required for 95% vs. 99% Confidence level, respectively, with a Margin of Error at 5% or less. Considering in this study that the hypothesized model was relatively simple with respect to the directional relationships between the variables, and there were eight identified construct variables with each being measured by three survey items, Table 1 guidance by Hair et al. (2010) suggested a more conservative approach with a sample of 500.

Table 1

Guidance for Minimum Sample Size Based on Model Complexity

Minimum Sample Size	Number of Constructs	Model Notes
100	1-5	3+ items for each construct (observed variables) and high item communalities of 0.6 or greater
150	1-7	Modest item communalities of 0.5 and no under identified constructs
300	1-7	Communalities of 0.45 or less and/or multiple under identified constructs
500	7 +	Some items may have low communalities and/or fewer than 3 measured items

Note. Adapted from *Determinants of Aviation Students' Intentions to Use Virtual Reality for Flight Training* [Ph.D. dissertation, Embry-Riddle Aeronautical University], by G.S. Fussell, 2020, Scholarly Commons (https://commons.erau.edu/edt/542/). Original guidance from *Multivariate Data Analysis* (p. 574), by J.F. Hair, W.C. Black, B.J. Babin, and R.E. Anderson, 2010, Prentice Hall.

The Westland (2010) formula provides a mathematical approach for calculating the minimum sample size required for SEM analysis and was the preferred method for this study because it does not require a population size. The formula accounts for the anticipated effect size, desired statistical power, probability level, and the number of latent and observed variables. There is precedented use of Westland's (2010) formula in studies with SEM analysis, which justifies its use for this study. The formula is as follows (Fussell, 2020):

$$n = \frac{1}{2H}(A\left(\frac{\pi}{6} - B + D\right) + H$$

$$+\sqrt{(A\left(\frac{\pi}{6}\right) - B + H)^2 + 4AH(\frac{\pi}{6} + \sqrt{A} + 2AB - C - 2D))}$$

Where:

$$A = 1 - \rho^2$$

$$B = \rho \arcsin(\frac{\rho}{2})$$

$$C = \rho \arcsin(\rho)$$

$$D = A/\sqrt{3 - A}$$

$$H = (\frac{\delta}{z_{1-\frac{\alpha}{2}} - z_{1-\beta}})^2$$

Using the online tool developed by Soper (2021) to calculate the required sample size based on the formula above, the minimum sample size for the study was set at 444 survey responses, considering the study included eight latent variables and 24 observed variables (three per construct). This mathematical estimation was also based on an anticipated effect size of 0.2, desired power level of 0.8, and a probability of 0.05, as recommended by the Soper (2021) calculator for SEM analysis.

The required sample size was planned to be recalculated if more data became available, using the Raosoft (2004) equations, to validate the success criteria. For practical reasons, the plan was to collect enrollment data from the FAA approved school websites to the extent possible and inquire about the number of students enrolled in the aircraft maintenance program when contacting each school to obtain approval for the study. Since this data was not made available, for the purposes of the study, a sample size of 444 was assumed to be adequate and appropriate to achieve priori thematic saturation/data saturation.

Sources of the Data

The source of data for the proposed study was from a cross-sectional survey that was used to collect quantitative data utilizing a five-point Likert Scale. The survey was administered digitally via email and social media to students enrolled in an aircraft maintenance program at an FAA-approved school and as permitted by the school. There were certain factors that prevented or made it difficult to conduct the survey in-person, such as temporary restrictions due to a widespread virus - COVID. Therefore, the online survey was the best option, though it may not have produced as many responses as desired. Many students may have been more hesitant to participate from a home environment (assuming many of the schools were practicing distance/hybrid learning more so than ever before), or some students may have had connectivity issues (e.g., no 24/7 access to internet). The data collection began as soon as the IRB application was approved (see Appendix), and the survey instrument was ready. The survey remained open for three months to give adequate time for students to participate and to increase the chances of achieving the required sample size. Whenever possible, students were sent a reminder email to take the survey.

Data Collection Device

An online survey instrument was used to collect quantitative data. There was a briefing/introduction section, to communicate the purpose of the study, a consent form with disclosure of participant rights, a short presentation for basic overview of VR as applied to maintenance training, and a set of questions to confirm that the student qualified to participate based on self-reported enrollment status (currently enrolled student) and age (an adult – 18 years of age or older). The consent form included information on the raffle (as incentive for participation) and raffle entry limitations. The survey then began with demographic information, such as age, gender, race, education status, current/previous VR use, followed by questions to measure the variables being studied.

The survey instrument consisted of 29 questions for measuring the constructs, three questions per each variable/construct being studied (24 questions), plus five demographic questions. Each of the 24 questions had a corresponding open response/text field as an opportunity for participants to provide an explanation for their responses, as to why they agree/disagree with a given statement. The total number of questions is mainly driven by the minimum number of items required to measure each construct with greater reliability and validity, as suggested by previous studies and with consideration for achieving higher confidence levels and lower margins of error in data analysis and results. Three measurement items per each construct were also inferred to be adequate based on the sample size guidance provided by Hair et al. (2010). The five demographic questions were selected based on research interests and were consistent with the demographic factors measured in similar studies on VR technology adoption. The intent of the demographic questions was to verify proportionate representation if any group is over/under-represented. The demographic questions and subgroup categories were as follows:

- Age
- Gender (Female/Male/Other/Prefer not to say)
- Race (Latino or Hispanic/African American/Asian/Caucasian/Native American/Other (specify)/Unknown/Prefer not to say)
- Education Status (First year/Second year/Third year or beyond)
- VR Experience Level (Never used VR/Used VR before, but not a frequent or current user/Frequent or routine VR user)

The predictive variables, as indicated in the TAM model, are perceived usefulness and perceived ease of use. In addition to these variables, five other relevant external variables were selected based on the literature review, to study external factors that influence perceived usefulness and perceived ease of use as determinants of behavioral intention of using VR for maintenance training. Operational definitions for these variables are provided in Figure 6. A five-point Likert scale was used to collect quantitative data, with 1 indicating *strongly disagree*, and 5 indicating *strongly agree* responses. Figure 6 also presents the survey questions, developed based on literature review and validated instruments from previous studies using TAM as the theoretical framework. The survey questions were modified to fit the technology being studied and its application and were refined/finalized with help from subject matter experts, as necessary, prior to administering the survey. The modification of the questions was justified by the fact that the theoretical model of choice in this study, TAM, is a technology independent model. This means TAM provides a theoretical framework that can be applied to any technology, but customization of the questionnaire is needed to add context of use for a given technology. Since its development, TAM has been implemented within different contexts as established through the questionnaire and consistently demonstrated to be valid in many previous studies (Lala, 2014).

Figure 6

Preliminary Survey Questions

Construct Variables	Operational Definition	Reference(s) for Definitions	Survey Questions/ Indicators	Reference(s) for Questions
Behavioral Intention (BU)	Degree to which an aircraft mechanic student intends to adopt VR technology for aircraft maintenance training.	Huang & Liaw (2018)	BI1: If made available, I will use VR for aircraft mechanic training.	Gong, et al. (2004); Park (2009); Sagnier et al. (2020)
			BI2: I intend to use VR for aircraft mechanic training if offered as an option in the future.	
			BI3: If made available, I would use VR for aircraft mechanic training regularly.	
Perceived Usefulness (PU)	Degree to which an aircraft mechanic student believes that using VR technology would be beneficial to his/her training.	Huang & Liaw (2018); Park (2009)	PU1: Aircraft maintenance training using VR would enhance my learning.	Park (2009); Sagnier et al. (2020)
			PU2: Aircraft maintenance training using VR will be useful for real world aircraft maintenance.	
			PU3: Using VR will make aircraft maintenance training more efficient.	
Perceived Ease of Use (PEU)	Degree to which an aircraft mechanic student believes that using VR technology for aircraft maintenance training would be effortless.	Huang & Liaw (2018); Park (2009)	PEU1: Learning to use VR for aircraft maintenance training will be easy for me.	Park (2009); Sagnier et al. (2020)
			PEU2: It will be easy to gain skills for aircraft maintenance using VR.	
			PEU3: I find VR easy to use for aircraft maintenance training.	
Self-Efficacy (SE)	Degree to an aircraft mechanic student has confidence that he/she is able to operate VR technology for aircraft maintenance training.	Huang & Liaw (2018); Levy & Green (2009)	PSE1: I feel confident in my ability to use VR for aircraft maintenance training.	Gong, et al. (2004); Levy & Green (2009); Park (2009)
			PSE2: I could use VR for aircraft maintenance training if someone showed me how to do it first.	
			PSE3: I feel confident in my ability to use VR for aircraft maintenance training if I see someone else using it before trying it myself.	
Perceived Health Risk (PHR)	Degree to which an aircraft mechanic student perceives the use of VR technology for maintenance training to be posing a health risk.	Tang et al. (2019)	PHR1: Using VR for aircraft maintenance training may negatively affect my health.	Tang et al. (2019)
			PHR2: I am concerned that use of VR for aircraft maintenance training may have an adverse outcome on my physical health.	
			PHR3: Using VR for aircraft maintenance training is not safe.	
Perceived Enjoyment (PE)	The degree to which using VR for aircraft maintenance training "is perceived to be enjoyable in its own right apart from any	Fussell (2020)	PE1: Using VR for aircraft maintenance training would be enjoyable.	Fussell (2020)
			PE2: Using VR for aircraft maintenance training would be exciting.	

	performance consequences that may be anticipated."		PE3: I would enjoy using immersive simulation technology such as VR for aircraft maintenance training.	
Performance Expectancy (PEXP)	The degree to which a student believes that using VR for aircraft maintenance training will improve his/her performance as compared to traditional learning methods	Fussell (2020); Lewis et al. (2013)	PEXP1: Using VR for aircraft maintenance training is more productive and efficient compared to traditional learning methods (textbooks, lectures, videos, etc.)	Fussell (2020)
			PEXP2: Using VR for aircraft maintenance training will improve my aircraft maintenance skills.	
			PEXP3: Using VR for aircraft maintenance training will improve the progression of my training.	
Perceived Behavioral Control (PBC)	The extent to which an aviation student feels able to control using VR technology for aircraft maintenance training	Fussell (2020)	PBC1: I could use VR for aircraft maintenance training if I could call someone for help if I got stuck.	Fussell (2020)
			PBC2: I could use VR for aircraft maintenance training if I had virtual instructor guiding me.	
			PBC3: I could use VR for aircraft maintenance training if I had only the manuals for reference.	

Note. The table presents the construct variables with their operational definitions and the corresponding survey questions, with corresponding sources.

Reliability and Validity

The validity and reliability of the study were addressed through the theoretical framework used, sampling method, survey instrument, and analysis applied to the data. The overall objective was to ensure that the results were accurate and meaningful. Use of the Technology Acceptance Model (TAM) establishes construct validity as the model has been demonstrated to be highly valid and successful in operationalizing the conceptual variables through survey methodology. A five-point Likert scale survey as the instrument of choice is common across studies on technology adoption. The preliminary survey questions were formulated based on the literature review and were further reviewed by research advisors. Furthermore, the survey questions/instrument were refined, as necessary, based on the results of the pilot study to ensure that the questions accurately measured each construct (reliability) and produced meaningful results (validity).

Full SEM also addresses convergent and discriminant validity for different measures of the same construct and uniqueness of the constructs, respectively. Composite Reliability (CR) values are used to determine if an observed variable should be removed from the model. Variables with CR values of greater than 0.7 were deemed acceptable and were retained in the model. Cronbach's alpha is another method of assessing instrument reliability and was reserved as an alternative technique for variables that showed low CR values. Use of CFA analysis prior to SEM also ensures that measures/constructs are valid and reliable, providing an opportunity to assess and adjust the hypothesized model and address assumptions required for SEM. Construct and convergent validity were addressed based on values for factor loadings and average variance extracted (AVE) from CFA output. AVE values of greater than 0.5 were deemed acceptable. The AVE value was used to assess the distinctiveness of each construct - discriminant validity. If the AVE value for a given factor is greater compared to the maximum shared variance (MSV) of the corresponding factor, then discriminant validity is deemed acceptable (Fussell, 2020). The selected statistical analysis was deemed to be the most appropriate for hypothesis testing based on ground theory as it applies regression equations to the data and has the potential to reveal new relationships among construct variables; therefore, the method provides the statistical validity needed to draw conclusions.

In terms of external validity, while effort was made to achieve diversity of student participants, the sample was limited to those schools that were willing to support the study. A survey methodology is highly based on self-reports, which some researchers argue is biased; however, there are no known methods that are 100% reliable and objective to measure the conceptual variables being studied.

Treatment of the Data

The data collected from the survey needed to meet certain assumptions to qualify for SEM analysis. SEM can only be applied to normally distributed and complete data. The assessment for normality was done through histograms and descriptive statistics, looking for any indication of kurtosis/skewness. In case the assumption was not met, the data was to be transformed using the SPSS software through one of a few methods. Any missing data needed to be addressed through an imputation technique. Missing values were identified in SPSS prior to conducting CFA. Depending on the situation, it may have been necessary to address the missing data by regression imputation, deleting the entire set of responses for a given participant(s) or substituting it with a value that is derived from the rest of the data (e.g., mean or case substitution). Likewise, any outlier data needed to be assessed to determine if it is best to remove, keep, or transform it (Truong, 2018).

Descriptive Statistics

Descriptive statistical analysis was conducted to test the data collected on TAM variables against the assumption of SEM and determine if data transformation was needed. More specifically, the descriptive statistics were used to reveal if the data was normally distributed vs. kurtotic and/or skewed. Normally distributed data means that values at extreme ends of the data set do not have a significant impact on the mean value, and the data distribution is bell-shaped, where 68.2% of observations lie between + or - 1

Standard Deviation (SD), 95.4% lie between + or - 2 SD, and 99.7% lie between + or - 3 SD from the mean (Anaesth, 2019).

The normality of the distribution is particularly important in achieving meaningful statistical analysis results. This is because the mean is used to calculate statistical significance (*P* value). If the mean turns out to be not representative of the data, then it would be inappropriate to use the mean value for inferential statistics as it could lead to incorrect interpretation of the results. While the central limit theorem suggests violation of the normality may not be an issue for a sample size of over 100 (Anaesth, 2019), it is best to test the data against the assumption, and if not met, transform the data prior to applying SEM analysis. This best ensures that the results are meaningful. For the purposes of this study, the normality of the data was tested using the Kolmogorov-Smirnov method in SPSS, which is appropriate for a large sample size. For the pilot study, the Shapiro-Wilk test was considered to be more appropriate due to a smaller sample size (Anaesth, 2019).

In summary, descriptive statistics was conducted on the demographic data collected and for each of the variables. The demographic data included the age, gender, and experience level with VR or aircraft maintenance or gaming. The mean values with SD for each of the variables provides some indication of participants' attitude toward VR, if there is general agreement/disagreement across all participants with respect to their perception of the technology, or if participant responses are neutral in some cases. Descriptive statistics results are provided in the form of a table/chart, values for mean/mode/median, standard deviation, etc., as appropriate (see Chapter IV).

Hypothesis Testing

The small-scale study could present a need to adjust the model and hypotheses; therefore, the hypothesis testing was planned to take place through SEM, first by testing for good model fit through CFA, then by testing the strength or significance of any relationships between the variables through SEM. Model fit indices of CFI, GFI, AGFI, NFI, CMIN/df, and RMSEA were used to draw conclusions on relationships between variables. Each of these indices has an acceptable range that the actual value must fall within to conclude that a relationship exists. The acceptable values for the model fit indices are presented in Table 2. Standardized regression weight and CR values were used for hypothesis testing and reporting. These values helped determine the direction of a given relationship between variables and significance of that relationship (Truong, 2018).

Table 2

Model Fit Summary from SEM

Model Fit Indices	Acceptable Range
CFI	> 0.93
GFI	> 0.9
AGFI	> 0.9
NFI	> 0.9
CMIN/df	3 or less
RMSEA	< 0.06

Summary

Chapter III provided details of the study methodology. The research approach was described to be quantitative, with the data being collected using a survey instrument. The chapter also described the challenges associated with estimating the population size and how the success criteria was set for the required sample size. The analysis technique and tools were also presented with a step-by-step description of how SEM was planned to be conducted. Finally, the chapter addressed the reliability and validity of the results based on the selected statistical analysis technique and applicable assumptions, factor loadings, and average variances.

Chapter IV will cover the actual results from the study, detailing the measures obtained from descriptive and inferential statistical analysis, to include the model fit indices. Chapter V will provide a discussion of the results and interpretations, along with implications and recommendations for future research.

Chapter IV: Results

This study attempted to address acceptance of VR technology by aircraft mechanic students through a Likert scale survey questionnaire that was designed to measure students' Perceived Usefulness (PU), Perceived Ease of Use (PEU), and Behavioral Intention (BI), as the construct variables supported by Technology Acceptance Model (TAM). Based on the literature review, the study also considered the external variables of Self-Efficacy (SE), Perceived Enjoyment (PE), Perceived Health Risk (PHR), Performance Expectancy (PEXP), and Perceived Behavioral Control (PBC) as contributors to PU and PEU. A model was developed to hypothesize the relationships between these variables and to determine the extent to which the variables affect the students' intent to use VR technology for aircraft maintenance training.

Chapter IV presents the results of the study in a sequential fashion, for all phases of the research process and analysis. First, the chapter covers findings from the Structural Equation Modeling (SEM) analysis of the pilot study data, explaining the outcome and providing a rationale for the final study decisions. Next, the chapter presents the results from the demographic data, descriptive statistics, reliability, and validity testing. Sample size for the final study analysis is also addressed as justification for switching from SEM to Exploratory Factor Analysis (EFA) for the final study analysis. Finally, the results from EFA for the final study are presented, addressing the hypothesized factors and the updated research question.

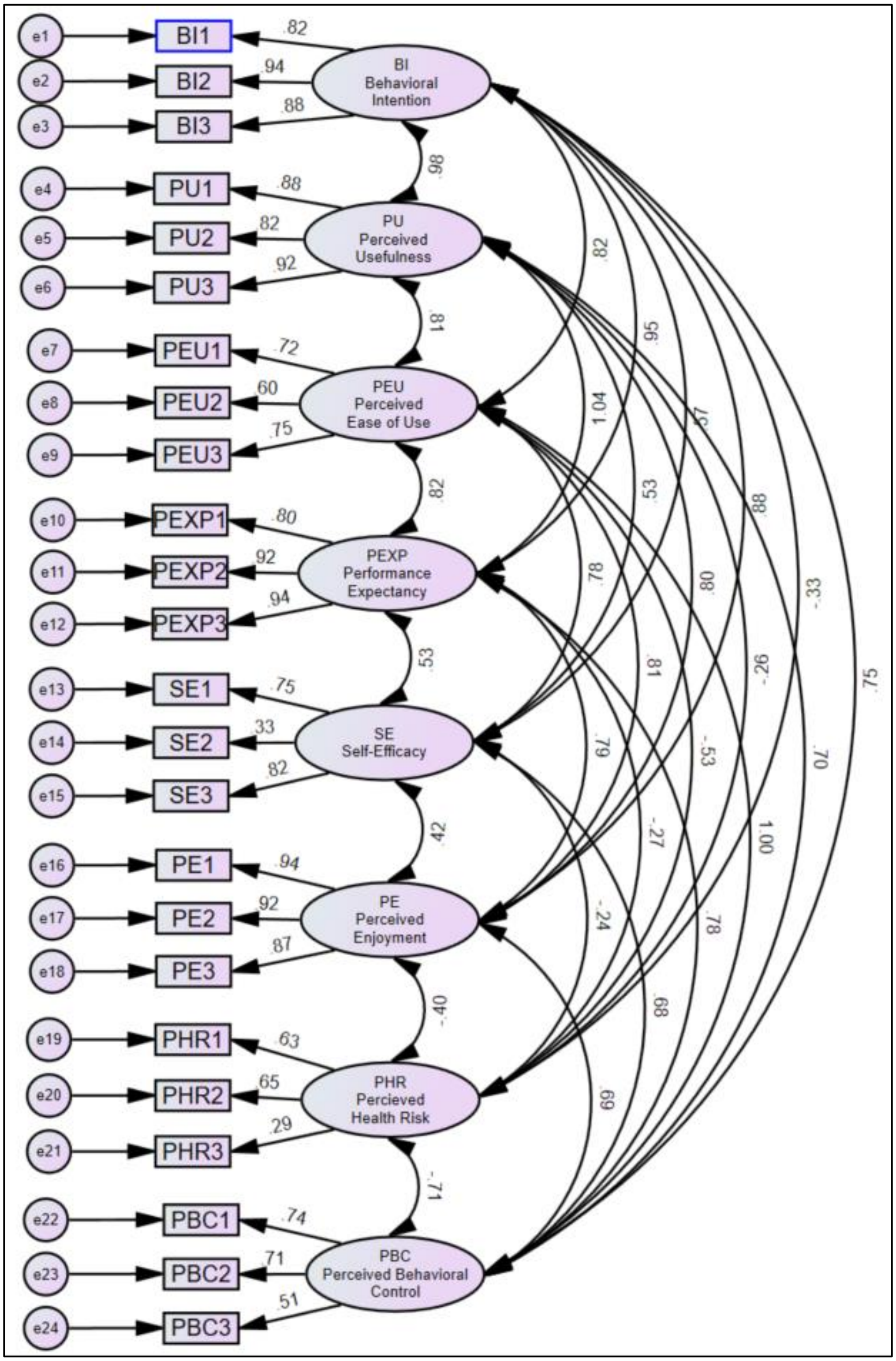
Pilot Study Results

A pilot study was conducted with aircraft mechanic students from Embry-Riddle Aeronautical University (ERAU). The survey was administered through Google Forms.

Students were invited to participate via email and in-person solicitation. A sample size of 55 participants was achieved. The data was prepared for CFA analysis as part of the full SEM process. First, AMOS was used to create a diagram with all latent and observed variables (indicators) and all possible relationships. The diagram is presented in Figure 7, along with CFA calculation outputs. The Notes for Model showed that the model is adequately over-identified, where the number of data points are higher than the number of parameters, with positive Degrees of Freedom. This indicated an ability to reject the model; hence, there was no need to fix the regression weights/factor loadings for the relationships between variables. With CFA being more sensitive to kurtosis than skewness, a kurtosis value of 5 and skewness value of 3 were used as thresholds (Truong, 2022). Normality was found to be acceptable based on the AMOS data. The highest values were 3.173 for kurtosis and 1.843 for skewness. There were no outliers found based on the Mahalanobis D-square values being less than 100 and acceptable (Truong, 2022).

Figure 7

First Specified CFA Model



The objective of the CFA analysis was to obtain an acceptable model fit to ensure validity and reliability of the measures and the constructs prior to running SEM. The process is iterative in nature, requiring removal of problematic variables and retesting. As an initial step, regression weights (otherwise known as factor loadings) were reviewed for all observed variables, where values above .5 were deemed acceptable. Factor loadings for PHR3 and SE2 showed unacceptable factor loadings. It is important to note that factor loadings alone are not sufficient for drawing conclusions on model fit. Model fit indices also showed poor model fit based on CFI, GFI, AGFI, NFI RMSEA values obtained (see Figure 8).

Figure 8

Model Fit Indices Results

Model Fit Indices	**Thresholds & Expectations**	**Values**	**Results**
CMIN/df (normed Chi-square)	<=3	**1.664**	Satisfactory
CFI	>=0.93	**.867**	Not Satisfactory
GFI	>=0.9	**.652**	Not Satisfactory
AGFI	>=0.9	**.534**	Not Satisfactory
NFI	>=0.9	**.733**	Not Satisfactory
RMSEA	<=0.06	**.111**	Not Satisfactory

Since good model fit was not achieved, a post-hoc analysis (model re-specification) was conducted. The purpose of the post-hoc analysis was to explore, revise the model by making necessary adjustments based on Modification Indices (MI) values, and retest the model. This was done in an effort to achieve a good model fit without convoluting or over-manipulating the model to fit the data. After reviewing the MI values for possible covariances between two error-terms or cross loading (regression between an item and a factor) and adjusting the model with one change at a time, a good model fit was still not achieved. The assessment also demonstrated poor construct reliability with multiple variables showing Cronbach's Alpha values below 0.7. Convergent and discriminant validity based on AVE and MSV values were also poor. Factor loadings, validity, and reliability results, focusing on out-of-range values, are shown in Table 3.

Table 3

Factor Loadings, Construct Reliability, Convergent & Discriminant Validity

Construct	Item Question	Factor Loading (≥0.5)	CR(≥ 0.7)	Cronbach's Alpha (≥0.5)	AVE (≥0.5)	MSV
BI	BI1	0.81				
	BI2	0.94	0.87		0.77	0.96
	BI3	0.88				
PU	PU1	0.88				
	PU2	0.82	0.87		0.76	1.08
	PU3	0.92				
PEXP	PEXP1	0.80				
	PEXP2	0.92	0.89		0.79	1.08
	PEXP3	0.94				
PE	PE3	0.87				
	PE2	0.93	0.92		0.83	0.78
	PE1	0.94				
PEU	PEU3	0.75				
	PEU2	0.60	0.72		0.48	1.01
	PEU1	0.72				
SE	SE1	0.75				
	SE2	0.33	0.65	0.59	0.45	0.60
	SE3	0.81				
PHR	PHR1	0.63				
	PHR2	0.65	0.51	0.45	0.30	0.51
	PHR3	0.29				
PBC	PBC3	0.51				
	PBC2	0.71	0.68	0.67	0.30	0.51
	PBC1	0.74				

It is important to note that MSV values are based on the correlations between the constructs. The initial analysis produced multiple correlations that exceeded the parameter of -1 to +1, as suggested by the unacceptable MSV values of > 1 (see Table 3). The out of bound values are referred to as Heywood cases and may be indicative of the sample size being too small for CFA analysis, the data being not normally distributed, or having problematic outliers (American Psychological Association [APA], 2018).

Attempts to fix the Heywood cases, by setting the parameter to 1 and reanalyzing the data, failed to achieve a good model fit or show improvements in convergent and discriminant validity. Running the risk of convoluting the model, and after revisiting the literature and consulting with subject matter experts (SMEs) on the results obtained with the several attempts of adjusting the model, it was concluded that the poor model fit, validity, and reliability results were potentially due to the small sample size. Therefore, the overall approach for the final study was revisited with respect to the number of variables hypothesized and the associated survey questions. Given the anticipated challenges with achieving the required sample size for full SEM analysis of the final study, the scope of the final study was strategically reduced by removing the external variables, as coordinated with SMEs. The survey instrument was updated accordingly to reflect only three of the main construct variables from TAM: PU, PEU, and BI.

Final Study Results

Demographics

The survey produced 55 student responses from the pilot study and 65 from the final. The responses from the pilot study were included as part of the full sample for analysis. Whenever possible, demographic data was analyzed to make comparisons and ensure the sample was representative of the population. Demographic data collected included participants' age, gender, race, education status, and experience level with VR technology. The results are presented in Table 4.

Table 4

Demographic Attributes of Participants

Attribute	Category	Frequency (*N*=120)	Percentage
Gender	Female	19	15.8
	Male	94	78.3
	Other/Prefer not to say/No Response	7	5.8
Race	Latino or Hispanic	18	15.0
	African American	9	7.5
	Asian	6	5.0
	Caucasian	68	56.7
	Native American	2	1.7
	Other/Mixed/Prefer not to say/No response	17	14.2
Education Status in AMT Program	1st year	36	30.0
	2nd year	51	42.5
	3rd year or beyond	28	23.3
	No response	5	4.2
VR Experience	Never used	15	12.5
	Used before, but not a frequent or current user	76	63.3
	Frequent or routine user	24	20.0
	No response	5	4.2

There was no published demographic information by the FAA approved AMT schools, and no demographic information was found on the age distribution of aircraft mechanic student population as a whole. Therefore, no comparison was made for participants' age. Participants' ages ranged from 18 years old to 67 years old, averaging around 25. Gender distribution was relatively consistent with a study conducted by Wang et al. (2016) as well as with the Aviation Technician Education Council (ATEC) Pipeline Report (2023) in that majority of students were male. Participant race was also close to the ATEC Pipeline Report (2023), with the majority of students (56.7%) being Caucasian and 29.2% minority, per the FAA definition of minority groups (FAA, n.d.-b). The ATEC Pipeline Report (2023) presented data on the current workforce of aircraft mechanics. It was assumed that the demographics for student aircraft mechanics would be similar. There was no data to compare to education status/enrollment distribution per the number of years in the program and VR experience level of the student aircraft mechanic population. Based on the data comparison for gender and race, the sample was deemed representative of the aircraft mechanic population.

Descriptive Statistics

Descriptive statistics of data was conducted in SPSS on the nine observed variables that were theorized to represent the constructs of Behavioral Intention (BI), Perceived Usefulness (PU), and Perceived Ease of Use (PEU). Table 5 lists the observed variables and associated survey questions. Participant responses were measured on a five-point Likert scale ranging from *strongly disagree* (1) to *strongly agree* (5). Table 6 shows the results from descriptive statistics, which includes the Mean, Standard Deviation, Skewness and Kurtosis.

Table 5

Final Survey Constructs, Observed Variables, & Survey Questions

Construct (data type)	Observed Variable/ Question Item	Survey Question
BI	BI1	If made available, I will use VR for aircraft maintenance training.
(metric)	BI2	I intend to use VR for aircraft maintenance training if offered as an option in the future.
	BI3	If made available, I would use VR for aircraft maintenance training regularly.
PU	PU1	Aircraft maintenance training using VR would enhance my learning.
(metric)	PU2	Aircraft maintenance training using VR will be useful for real world aircraft maintenance.
	PU3	Using VR will make aircraft maintenance training more efficient.
PEU	PEU1	Learning to use VR for aircraft maintenance training will be easy for me.
(metric)	PEU2	It will be easy to gain skills for aircraft maintenance using VR.
	PEU3	I find VR easy to use for aircraft maintenance training.

Table 6

Descriptive Statistics Results

Variables	N	Mean	Std. Deviation	Skewness	Std. Error of Skewness	Kurtosis	Std. Error of Kurtosis
BI1	120	4.17	1.103	-1.332	.221	1.079	.438
BI2	120	4.07	1.262	-1.352	.221	.792	.438
BI3	120	3.76	1.283	-.822	.221	-.270	.438
PU1	120	4.09	1.115	-1.293	.221	1.051	.438
PU2	120	3.83	1.266	-.970	.221	-.031	.438
PU3	120	3.86	1.183	-.928	.221	.111	.438
PEU1	120	4.11	1.067	-1.189	.221	.985	.438
PEU2	120	3.48	1.174	-.419	.221	-.503	.438
PEU3	120	3.60	1.111	-.238	.221	-.725	.438

Normality. Descriptive statistics confirmed no missing values but also indicated that the data is not normally distributed. All nine variables proved to be negatively skewed with the majority of the responses being above the average values shown in Table 6. For BI1 and BI2, the majority of the participants expressed agreement/strong agreement that they would use VR for aircraft maintenance training if made available or if offered as an option in the future. Similarly, the majority of participants responded *agree* or *strongly agree* on PU1 and PEU1, that aircraft maintenance training using VR would enhance their learning, and learning to use VR for aircraft maintenance training will be easy for them. For variables BI3, PU2, PU3, PEU2, and PEU3, the Mean values were below 4 (*agree*), which means there were a higher number of neutral ratings by participants. Specifically, variables BI3, PU2, and PU3 showed average scores above 3.5, which is still higher than a neutral rating. This indicates general but not strong agreement by participants that 1) If made available, they would use VR for aircraft maintenance training regularly (BI1); 2) VR training will be useful for real world aircraft maintenance

(PU2); and 3) Using VR will make aircraft maintenance training more efficient (PU3). PEU2 and PEU3 had the lowest averages and standard deviation (SD), M = 3.48 (SD = -0.419) and M = 3.60 (SD = -0.238), respectively. This means that all responses clustered around the average score. It also indicates that although many participants responded that it will be easy to gain skills for aircraft maintenance using VR (PEU2), and they find VR easy to use (PEU3), they are less certain about the overall ease of use of VR or are anticipating some challenges.

Skewness and kurtosis were also examined with Kolmogorov-Smirnov and Shapiro-Wilk test results. All variables showed significance values of <.0.05 for both tests, which indicated a normality issue. However, considering these tests are sensitive to sample size, where a larger sample size may show significant values (Truong, 2022), histograms were examined next to make conclusions on normality. Variables PEU2 and PEU3 were the closest to being normally distributed. All other variables had normality issues, which suggested that the most appropriate EFA extraction method would be the Principal Axis Factoring (PAF) (Hui, 2021).

Outliers. Outlier data was examined as indicated on the box plots and also assessed through the Mahalanobis Distance technique. While the box plots showed potential outliers, none of the cases were identified as extreme outliers. To further examine the data, the SPSS linear regression function was used to obtain Mahalanobis values. Then the Mahalanobis D-square values were calculated in SPSS using the Cumulative Distribution Function for Chi Square (CDF.ChiSq) in the numerical expression. With all Mahalanobis D-square values being above .001, it was confirmed that there were no problematic outliers in the data (Truong, 2016).

Data Treatment. The data showed no missing values or outliers that would require an imputation. The descriptive analysis confirmed skewness and kurtosis (non-normal distribution). As such, several transformation techniques were attempted using one of the most skewed variables, BI1, to see if the data distribution could be improved for the purposes of using EFA techniques that require a normal distribution. The techniques attempted were Squared, Logarithm, Inverse, and Square Root. While the skewness and kurtosis values showed some improvement, a normal distribution was not achieved. This was clearly evident by observing the histograms. Hence, the transformation attempts substantiated the use of the Principal Axis Factoring (PAF) method for EFA extraction. It is important to note that with skewness and kurtosis tests being sensitive to the sample size, in many cases a normal distribution may not be achievable even after data transformation attempts, which should not stop researchers from continuing with multivariate analyses (Truong, 2022). In some cases, evidence of skewness is an indication of natural bias of opinions in a given population rather than an indication of a bias introduced through the research design. This means that the attitudes towards VR technology expressed by student aircraft mechanics may naturally be strongly skewed but still representative of the population (Pennsylvania State University, 2024).

EFA Assumptions

Prior to conducting EFA analysis, several tests were needed to ensure that EFA specific assumptions were met. Other than having metric data type, which the survey results satisfied, EFA requires a large sample size, homogeneity of the sample, and inter-correlation among variables. The adequacy of the sample size was determined based on the Measures of Sampling Adequacy (MSA) values from Anti-image analysis. All MSA values were above 0.5, which indicated an adequate sample size for individual variables (Truong, 2016). Inter-correlation was assessed using Kaiser-Meyer-Olkin (KMO) and Bartlett's tests. A KMO value of more than 0.5 and a significant Bartlett test result were considered acceptable. The test showed a KMO value of 0.932, and the Bartlett test of sphericity showed to be significant at less than .001 (below the threshold of .05). Finally, homogeneity was determined through the Levene test across the categorical variable of student's education level. Homogeneity is an indication of variance equality, when the variances in multiple groups are not significantly different. Hence, the test looked for low variances among the groups. Non-significant Levene test results with significance values of more than .05 were considered acceptable (Truong, 2016). In this case, all significance values were more than .05, which met the homogeneity assumption.

EFA Analysis Results

EFA was conducted using SPSS. EFA requires selecting extraction and rotation methods. The PAF extraction, which is a least-squares estimation technique without using any assumptions about the distribution, was the method of choice due to the data not being normally distributed (Hui, 2021). The first attempt of EFA analysis was based on a minimum Eigenvalue of 1 for factor extraction. Oblique rotation was deemed to be more appropriate as opposed to the orthogonal rotation method, to allow correlation between factors as suggested by the Technology Acceptance Model (TAM) for the theorized factors/constructs of Perceived Usefulness (PU), Perceived Ease of Use (PEU), and Behavioral Intention (BI). The oblique rotation method is also known to produce more accurate results (Hui, 2021) in testing the relationship between the extracted factors, though it is harder to interpret (Truong, 2016). In the initial EFA analysis with oblique rotation, coefficients with a value less than .5 were suppressed.

Based on the Eigenvalues, the SPSS output for Total Variance Explained showed that all observed variables belonged to a single factor. The single factor extracted accounted for 69.33% of the variance (see Figure 9). Hence, no rotation was possible.

Figure 9

Initial EFA Analysis Showing the Variance by a Single Factor

Total Variance Explained

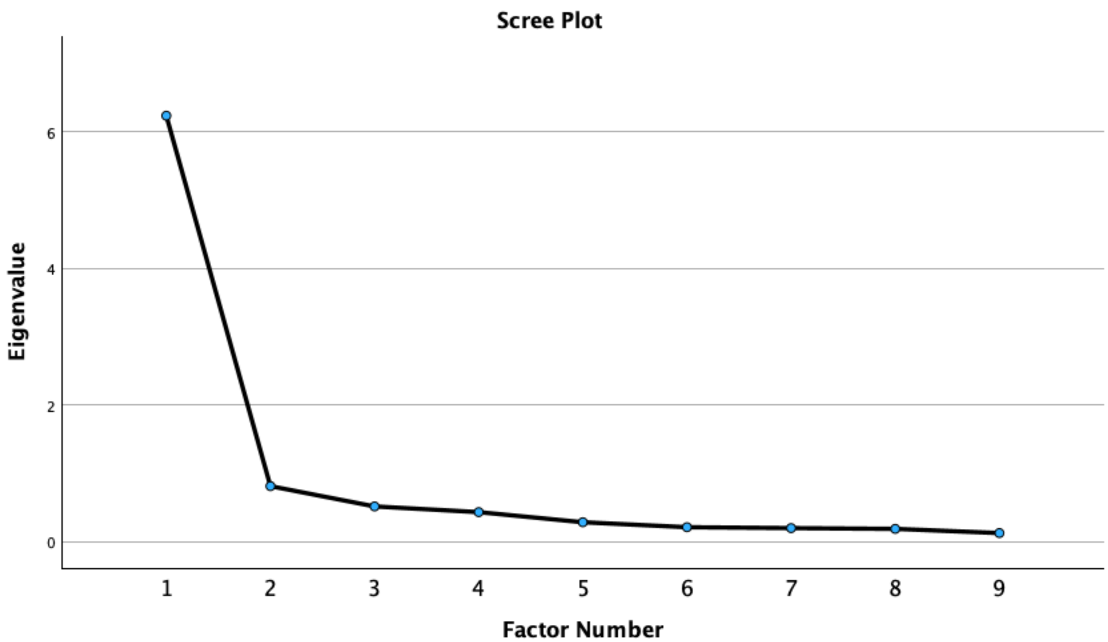
Factor	Initial Eigenvalues Total	Initial Eigenvalues % of Variance	Initial Eigenvalues Cumulative %	Extraction Sums of Squared Loadings Total	Extraction Sums of Squared Loadings % of Variance	Extraction Sums of Squared Loadings Cumulative %
1	6.239	69.326	69.326	5.939	65.994	65.994
2	.812	9.025	78.352			
3	.515	5.721	84.073			
4	.431	4.794	88.867			
5	.283	3.148	92.015			
6	.210	2.338	94.353			
7	.197	2.193	96.545			
8	.186	2.067	98.612			
9	.125	1.388	100.000			

Extraction Method: Principal Axis Factoring.

The scree plot was reviewed to see if potentially multiple factors could be extracted based on the Cartel's method (Hui, 2021). Looking at the bend points of the line to find the point at which the line changes slope, otherwise known as the elbow of the line, is how the scree plot was interpreted. In some cases, the interpretation can be subjective. For example, looking at the scree plot from the initial EFA analysis (see Figure 10), one could argue that the elbow of the line is at component number 3, which would mean at least two factors are extractable. It is important to note that only one factor on the scree plot had an Eigenvalue above 1. The Eigenvalue is representative of the total amount of variance in the data that can be explained by a given factor.

Figure 10

Initial EFA Analysis: Scree Plot Output



Given the exploratory nature of the study and the subjectiveness in interpreting the scree plot, the EFA analysis was repeated with a fixed number of three factors as a theory-driven approach and as suggested by TAM. The second EFA analysis was conducted by suppressing the low coefficient values below .3. The scree plot looked the same. The analysis produced three factors as commanded. Factor 1 included the following variables: BI1, BI2, BI3, PU1, PU2, PU3, and PEU2. Factor 2 included only two items: PEU1 and PEU3. Based on the set criteria for the analysis, Factor 3 did not show any items associated with it that would make it plausible. Typically, factors with at least three items are considered more explainable. This is especially important if the intent is to continue with CFA analysis as to avoid any identification issues (Hui, 2021). However, with CFA no longer being within the scope of this final study, Factor 2 with only two items was retained in the EFA analysis. Figure 11 shows the variances from the three-factor EFA analysis. Factor 1 accounted for 66.74% of the total variance, Factor 2 for 6.3%, and Factor 3 for 2.28%.

Figure 11

Final EFA Analysis Showing Variance Explained by Three Factor Extraction

Total Variance Explained

	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
Factor	Total	% of Variance	Cumulative %	Total
1	6.007	66.741	66.741	5.909
2	.567	6.297	73.038	3.863
3	.206	2.284	75.322	.206

Extraction Method: Principal Axis Factoring.

a. When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

The criteria applied by some researchers suggests that all factors considered should account for 70% to 80% of the total variance (University of California at Los Angeles [UCLA], n.d.). In this case, Factor 1 and Factor 2 together explained 73.02% of the variance, which also showed support for the decision to retain Factor 2 even though it included only two items (PEU1 and PEU3). Factor 3 was not considered as plausible due to its low contribution to total variance (2.28%), and the lack of any items being associated with it.

Pattern and structure matrices were used to interpret the results (see Figures 12 and 13). The pattern matrix indicates the factor loadings (partial standardized regression coefficients) for each item, while the structure matrix shows the correlation between the items and the factors (Hui, 2021).

Figure 12

Final EFA Analysis: Pattern Matrix Showing the Factor Loadings for Each Item

Pattern Matrix[a]

	Factor		
	1	2	3
BI1	.884		
BI2	.939		
BI3	.861		
PU1	.845		
PU2	.896		
PU3	.874		
PEU1		.752	
PEU2	.665		
PEU3		.830	

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 3 iterations.

As shown in Figure 12, higher factor loadings indicated a stronger influence by Factor 1 and Factor 2 on each associated item/variable. Items that appear under multiple factors present a possible cross-loading issue. This happens when a given item is not necessarily associated with a single factor, but rather multiple factors contribute to the variance of that item (UCLA, n.d.) In this case, the pattern matrix did not indicate any cross-loading items that would require deletion. For all items, the factor loadings were above .5, the acceptable threshold (Truong, 2016). It is important to note that contrary to what was anticipated, PEU2 showed an association with Factor 1, which included the BI and PU items, instead of Factor 2, which included PEU1 and PEU3. This means that the participant responses to PEU2, "It will be easy to gain skills for aircraft maintenance using VR", did not support the idea of PEU2 being associated with the construct of perceived ease of use, as theorized. PEU1 and PEU3 were shown to be the only two measures of the PEU construct. However, the relationship between Factor 1 and PEU2 was also not as strong as the relationships between Factor 1 and other items (BI and PU). This is a possible indication of the PEU2 survey question being ambiguous. Additionally, the results showing all BI and PU items under the same factor (measuring the same construct) could be an indication of perceived usefulness of VR being highly correlated with behavioral intention to use VR as suggested by TAM. The results showed no particular distinction between the two (BI and PU) from a statistical perspective, which also means participants responded to the associated survey items in a similar pattern.

The structure matrix supported the same results, in terms of which items are associated with which factor based on simple correlation values, as shown in Figure 13.

Figure 13

Final EFA Analysis: Structure Matrix Showing Item Correlations with Each Factor

Structure Matrix

	Factor		
	1	2	3
BI1	.865	.583	
BI2	.926	.626	
BI3	.891	.638	
PU1	.865	.605	
PU2	.901	.623	
PU3	.870	.592	
PEU1	.574	.797	
PEU2	.685	.488	
PEU3	.626	.863	

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

The structure matrix showed that PEU1 and PEU3 are strongly correlated with Factor 2 with values of .8 and .86, respectively. The remaining factors were highly correlated with Factor 1. The weakest correlation with Factor 2 was PEU2 as anticipated after observing the pattern matrix. Still, the correlation of PEU2 with Factor 1 was .69, above moderate levels. Variables BI1, BI2, BI3, PU1, PU2, PU3 showed a strong correlation with Factor 1 with values above .7.

Theoretically, both the pattern and the structure matrix confirm that based on the participant responses, PEU1 and PEU3 measured the same construct (Factor 2), whereas all other items measured a different construct (Factor 1). From a practical perspective, the results show that participants responded in a consistent way to all survey items associated with BI, PU, and specifically PEU2. The findings contradict what was theorized or expected -- three separate constructs/factors that would represent the participants' behavioral intention of using VR technology for aircraft maintenance training, perceived usefulness, and perceived ease of use of VR. This means that while Factor 2 can remain named as "perceived ease of use", Factor 1 would require additional considerations by revisiting the actual survey items and possibly any corresponding comments, to produce a representative name. This also means that the results did not fully support the model structure suggested by TAM. The implications of these findings are further discussed in Chapter V.

The factor correlation matrix was used to examine the correlation between Factor 1 and Factor 2 (see Figure 14). The correlation coefficient value was .69, which substantiated the use of the Oblique rotation technique for EFA.

Figure 14

Final EFA Analysis: Factor-to-Factor Correlation Matrix

Factor Correlation Matrix

Factor	1	2	3
1	1.000	.687	-.007
2	.687	1.000	.016
3	-.007	.016	1.000

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

Reliability and Validity

Construct Reliability. Construct reliability was assessed prior to conducting EFA analysis and after the EFA results were obtained. Cronbach's Alpha was used as one of the most popular methods (Truong, 2016). The test examines the degree of consistency between multiple measurements of the same factor/construct. Prior to EFA analysis, the test was conducted with the combination of measurements for the theorized constructs. Then, the test was repeated after the EFA to assess the new set of relationships as suggested by the analysis. The Cronbach's Alpha test results from SPSS were used to determine if any items should be removed from the analysis a. poor measure of a given construct. Items with Cronbach's Alpha values of .7 and greater were considered as high in reliability and measuring the same construct. Inter-item correlation threshold was set at .3 or higher as an acceptance criterion. Similarly, the correlated item total correlation was expected to be above .3 (Truong, 2016).

For Behavior Intention, measured by items BI1, BI2, and BI3, Cronbach’s Alpha was .923. The inter-item correlation and corrected item-total correlation values were acceptable. Perceived Usefulness, measured by items PU1, PU2, and PU3 also showed good construct reliability with Cronbach's Alpha value of .913 and acceptable correlations. Similarly, the test for Perceived Ease of Use measured by items PEU1, PEU2, and PEU3 produced satisfactory results with Cronbach's Alpha value of .740 and correlation coefficient values above .3.

Cronbach's Alpha and correlations were also examined for the two factors/constructs suggested by the EFA results. For Factor 1, the measurements items tested included, BI1, BI2, BI3, PU1, PU2, PU3, and PEU2. For Factor 2, there were only two measurement items, PEU1 and PEU3. Both Factor 1 and Factor 2 showed high construct reliability with Cronbach's Alpha values of .950 and .769, respectively. The correlation coefficients were also acceptable. Results are presented in Table 7 and Figures 15 and 16.

Table 7

Construct Reliability Test Results: Cronbach's Alpha for Measurements of Factors 1 and 2

Factor/Construct	Measurement Items	Cronbach's Alpha (($\geq$ 0.7)
Factor 1	BI1 BI2 BI3 PU1 PU2 PU3 PEU2	.950
Factor 2	PEU1 PEU3	.769

Figure 15

Construct Reliability Test Results: Inter-Item Correlations for Factors 1 and 2

Inter-Item Correlation Matrix

	BI1	BI2	BI3	PU1	PU2	PU3	PEU2
BI1	1.000	.819	.771	.753	.761	.727	.600
BI2	.819	1.000	.825	.820	.801	.823	.608
BI3	.771	.825	1.000	.738	.808	.769	.630
PU1	.753	.820	.738	1.000	.779	.749	.569
PU2	.761	.801	.808	.779	1.000	.809	.654
PU3	.727	.823	.769	.749	.809	1.000	.582
PEU2	.600	.608	.630	.569	.654	.582	1.000

Inter-Item Correlation Matrix

	PEU1	PEU3
PEU1	1.000	.625
PEU3	.625	1.000

Figure 16

Construct Reliability Test Results: Corrected Item-Total Correlations for Factors 1 and 2

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
BI1	23.09	41.277	.839	.722	.942
BI2	23.19	38.711	.896	.830	.937
BI3	23.50	38.908	.863	.757	.940
PU1	23.17	41.199	.834	.722	.943
PU2	23.43	38.902	.879	.781	.939
PU3	23.40	40.276	.846	.744	.941
PEU2	23.78	42.697	.672	.468	.955

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
PEU1	3.60	1.234	.625	.391	.
PEU3	4.11	1.139	.625	.391	.

Validity. The discriminant validity of the study was addressed by examining the pattern and structure matrices produced by the EFA analysis. Discriminant validity relates to the accuracy of the measurement instruments, in other words, the extent to which factors/constructs are the different from each other (Hui, 2021). In Figure 12, the pattern matrix shows that measurement items load on a single factor/construct. The lack of cross-loadings of measurement items between Factor 1 and Factor 2 is an indication of acceptable validity. Discriminant validity was also confirmed through the factor correlation matrix (see Figure 14), with the correlation of Factor 1 and Factor 2 being .687, just below .7 as the threshold.

Summary

Chapter IV presented the pilot study and final study results, including the sample demographics, descriptive statistics, and EFA results. Confirmatory Factor Analysis (CFA), applied to the pilot study data failed to show a good model fit. For this reason, and as a result of the final study sample size being less than anticipated, the analysis applied to the final study was exploratory (EFA). The chapter presented evidence for the sample being representative of the target population based on the demographic data comparisons. Descriptive statistical results were presented to show that there were no issues with problematic outliers, though the data was not normally distributed. The results from assessing compliance with EFA assumptions were also presented, with respect to sample size adequacy, homogeneity, and inter-correlation among variables. Finally, Chapter IV presented the initial and final EFA analysis with the PAF extraction method and oblique rotation, also addressing the reliability and validity of the study. Chapter V will discuss the study results to provide conclusions and recommendations.

Chapter V: Discussion, Conclusions, and Recommendations

This study explored student aircraft mechanic survey data for theorized constructs of Perceived Usefulness (PU), Perceived Ease of Use (PEU) and Behavioral Intention (BI). These constructs were measured through observed variables (three survey items per construct) and expected to be correlated with each other based on the literature review and as suggested by TAM. The initial intent of the study was confirmatory. However, since the final study sample size was not adequate for SEM analysis and considering a good model fit was not achieved through CFA analysis of the pilot study data, the scope of the study was changed to exploratory. The data was analyzed through the EFA method. Hence, hypothesis testing was not possible. Instead, the objective of the final study was to detect underlying factors from the measurement items by using an interdependence multivariate statistical method, which does not discriminate between independent and dependent variables (Truong, 2016). The purpose of EFA was to extract the factors/constructs through data reduction. In other words, EFA helps discover the factor structure based on the interrelationships and commonality of the variables (Truong, 2016). Therefore, the research question that the final study attempted to address is, What are the underlying factors of the student aircraft mechanic survey? Chapter V provides a discussion of the final study results, the conclusions drawn based on those results along with considerations for theoretical/practical implications. The chapter also discusses encountered limitations and highlights recommendations for future research.

Discussion

Characteristics of Participants

The study analysis considered 120 cases of participant responses for exploration, of which 55 cases represented ERAU aircraft mechanic students -- responses from the pilot study. The demographic data offered some insight on participant characteristics. While the average age of participants was around 25 years old and the median age was 21, the age range was wider than anticipated with the oldest reported participant being 67 years of age. Potentially, 31 participants (25.83%) were 31 years of age or older (21 confirmed based on the responses). This presents considerations for generational differences between the student aircraft mechanics which could influence the perception of VR technology. The technology did not exist in the same capacity decades ago, and adoption of the current VR capabilities may vary from one generation to another.

Considering the current aircraft mechanic workforce demographics from the ATEC Pipeline Report (2023), the gender distribution was anticipated to be male dominant as the results suggested. In other words, it is safe to assume that the population of aircraft mechanic students in the U.S. has a majority of males since the workforce data also showed a similar distribution. Actual enrollment data was not available to compare and substantiate this assumption.

The race distribution was also found to be consistent with the ATEC Pipeline Report (2023) with the majority (56.7%) of participants being Caucasian and 29.2% a racial minority (Asian, Latino/Hispanic, African American, Native American) (FAA, n.d.-b). In further examining the individual responses, specifically the explanations provided by participants in open text format, there was some evidence that comprehension of the questions by some participants may have influenced the results. There were many grammatically poor comments, and comments that addressed the usefulness of VR when responding to a survey item that was intended to measure the perceived ease of use of VR for aircraft maintenance training. For example, for PEU2, "It will be easy to gain skills for aircraft maintenance using VR" the following comments provided implied that participants are potentially referring to the usefulness of VR rather than ease of use:

- *I don't think I could gain real life skills in VR; however, I gain more knowledge*.
- *My Pearnsol option* [sic] *on VR is that is could be very helpful on training and showing how tasks or projects could be completed. Also, if a person is struggling on doing a task they could see a video or picture on how to complete there* [sic] *task.*

One could confuse the concept of usefulness with ease of use when not understanding the context or intent of the question. However, the comments observed could be an indication of some participants not understanding or misinterpreting the question possibly due to a language barrier, especially if English is not their primary language. Alternatively, the observations could indicate that the question was too ambiguous and not truly measuring perceived ease of use. If the majority of participants treated the PEU2 survey item the same as usefulness of using VR for aircraft maintenance training, it could explain why the EFA results showed PEU2 being grouped with PU items under the same factor. For the purposes of being able to conduct the study online to achieve an adequate sample size and to address any risks associated with language as a barrier, it was assumed that all participants were proficient enough in

English language to understand the survey questions. This assumption is supported by the minimum requirements for entry into an FAA approved aircraft maintenance program within the U.S. as the instructions are conducted in the English language. Also, measures were taken in the consent form to avoid participation by students who are not residents of the U.S. Therefore, for this study it is more likely that the PEU2 survey item was not written clearly to encourage participants to respond to the ease of use of VR. The implications of this on actual adoption of VR technology for aircraft maintenance training may in fact be minimal because of the nature of this study being exploratory. However, there are theoretical implications that need to be considered with respect to confirmatory research and hypotheses testing for relationships between the various constructs suggested by a given theoretical framework for technology acceptance. For example, researchers should utilize all available data (quantitative and qualitative) to update and validate the survey instrument during a pilot study phase to make sure the survey items are measuring the constructs as intended. To rule out the possibility of language barriers/comprehension issues, researchers should also consider additional demographic questions to learn more about the participants' backgrounds. Measures can also be taken through the consent form to better screen participants for English proficiency. Establishing means to follow-up with participants when the responses do not seem to address the intent of the question, and/or analyzing the qualitative data to find an explanation for any unanticipated results/observations is also recommended.

The majority (63.3%) of the participants had some experience with VR but were not frequent or current users, whereas 20% were frequent/routine users of VR. Those who never used VR before or did not respond to the question accounted for 16.7% of all cases. There is no credible or scholarly data specifically for the student aircraft mechanic population with which to compare these results. However, VR technology is easily accessible and popular in the entertainment/gaming industry. Therefore, it is assumed that the majority of the student population within the U.S. are familiar with VR and have previously used it at least one time. Some support for this assumption is provided in a study with Georgia Tech college students by Soylu and Lee (2024), which showed that half of the students actually owned a VR headset.

The intent of the demographic data was to make comparisons and ensure that the sample represented the population of student aircraft mechanics within the U.S. While for participants' age, VR experience, and education/enrollment status there was no data with which to compare, the results were deemed as partially generalizable to the population of student aircraft mechanics based on the gender and race distribution. However, there may be generalizability limitations with respect to other demographics such as institution and age. The final study participants were not required to state which AMT school they are attending. Participant screening was addressed through the consent form only. The only confirmation of sample diversity was based on the volunteered data, with about 23% of participants identified as students from FAA approved AMT schools listed in Table 8.

Table 8

Final Study Sample Diversity Based on Volunteered Data

AMT School Name	AMT School U.S. State Location	Known Number of Participants
Everett Community College	Washington	2
Enterprise State Community College	Alabama	2
Purdue University	Indiana	1
Dutchess Community College	New York	1
Lake Superior College	Minnesota	1
Illinois State University	Illinois	3
Wayne Community College	North Carolina	3
Tarrant County College	Texas	1
Cape Cod Community College	Massachusetts	1

To further address generalizability limitations of the final study results from the perspective of participant demographics, a general comparison was made between the pilot study and final study participants. This comparison was done after the analysis and interpretation of the results based on SME recommendations to show improved generalizability. The objective was to also justify the inclusion of the pilot study data in the final analysis, considering the pilot study data was from ERAU students and made up about 46% of the sample analyzed. In this case, participant age and level of VR experience were of primary interest due to the lack of population data to compare with.

Both the pilot study sample (N=55) and final study sample (N=65) were deemed comparable on the basis of participant age and level of VR experience. The pilot study participants averaged around 22 years of age, while the final study participants averaged around 28 years of age. The difference could be attributed to one participant in the final study being 67 years of age and the nine participants who did not specify their age. However, the 67-year-old participant was not found to be a problematic outlier (through statistical testing) for the purposes of the final analysis. The median age for the pilot study group and final study group were closer, 21 years and 25 years, respectively. Finally, the mode for both groups was the same, 20 years of age. As for the level of VR experience, the majority of student participants (over 75%) in both the pilot study and final study indicated familiarity with VR (frequent or routine VR user, or used VR before, but not a frequent or current user). Additionally, the pilot study and final study participants seemed to be similar on the basis of gender (majority being male) and race (majority being Caucasian). This general comparison of the pilot study and final study groups substantiated the EFA decision to analyze both data sets together as they seem to be equal based on participant demographics, which in turn strengthened the generalizability of the EFA results and conclusions to the target population.

Discussion of the EFA Decisions & Results

The EFA analysis of the final study data required a series of decisions that are somewhat subjective and debatable amongst researchers. Main decision points included but are not limited to the data treatment (i.e., removing/keeping potentially problematic outliers or correcting skewness/kurtosis of the data distribution), extraction method, and rotation method.

Outliers. In this study, outlier data was assessed using the Mahalanobis Distance technique and by examining the box plots. The Mahalanobis Distance technique is a numerical method for outlier identification as it is based on the standard deviation and distance between two data points. While the Mahalanobis Distance technique did not show any outliers, the box plots suggested multiple outlier cases that could be considered as problematic, though none were identified as extreme outliers. Given the nature of this study, it is possible for outlier data to be beneficial (Truong, 2016) in a sense that it is expected to see some distinctly different ratings by student aircraft mechanics based on their subjective perceptions. In other words, when it comes to measuring responses through hypothetical questions on perceived usefulness of VR, perceived ease of use of VR, and behavior intention to use VR for aircraft maintenance training, each participant is entitled to their own opinion and contributes to the characteristics of the population. Having said that, outlier data that is not representative of the population could distort the analysis results (Truong, 2016). Considering the acceptable values obtained through the Mahalanobis Distance technique, and due to the lack of any obvious indication for the data being invalid, the outlier cases identified by the box plots were not examined any further. All cases were retained for the EFA analysis.

Normality. The data distribution was assessed for normality using histograms and skewness/kurtosis values. All observed variables (BI1, BI2, BI3, PU1, PU2, PU3, PEU1, PEU2, and PEU3) had a normality issue, many to great extent. While normality is preferred for multivariate statistical analysis, SPSS does offer an EFA extraction technique, the Principal Axis Factoring (PAF), suitable for a non-normal distribution (Hui, 2021). After trying various popular methods for transforming the data (i.e., log, inverse, squared, etc.) for one of the extremely skewed variables (BI1) and not being able to achieve normality, the PAF extraction method was deemed as the most appropriate. Other EFA extraction methods, such as the Principal Component Analysis (PCA) and Maximum Likelihood (ML) require a normal distribution (Hui, 2021). Again, EFA decisions based on the normality of the data are somewhat subjective. Normality tests are sensitive to the sample size. Some researchers may continue with other types of EFA extraction methods if normality is not achieved but is improved enough to be close to expected levels or simply assume normality based on a large sample size (Truong, 2016).

The normality assessment in this study does have some implications on the aircraft mechanic student population. At a minimum, it raises the question of the attitudes of the student population being naturally skewed or not. In this study, the sample was deemed as representative of the target population based on the gender and ethnicity demographic data only. This does not mean that there are no generalizability limitations with respect to other population characteristics, for example due to about 46% of participants being from ERAU. However, based on volunteered information, there was evidence that the sample did include AMT students from other states across the U.S., such as Washington, Alabama, Indiana, New York, Minnesota, Illinois, North Carolina, Texas, and Massachusetts. From a practical perspective, the selected EFA technique supported the analysis of the skewed data. However, the aviation industry could benefit from further exploring and understanding the characteristics of the AMT student population as a whole as well as in subpopulations of interest. For example, the size of the AMT school establishment may be an influential factor on student attitudes towards VR for aircraft maintenance training. Samples from these smaller establishments (subpopulations) may similarly produce skewed data, while the data from larger establishments may be more normally distributed. Understanding these differences could help industry practitioners make better decisions, from making inferences about the acceptance of VR for aircraft maintenance training based on the establishment size, to finding ways to implement the technology in a way that will serve its intended purpose and encourage students from all establishments (big and small) to use VR.

EFA Results. Often, the EFA ends up being an iterative process based on the decisions made every step of the way. For this study, EFA was conducted twice prior to making conclusions. Both attempts used the PAF extraction method and oblique rotation (Direct Oblimin). The rotation method was selected based on the expected correlation between the hypothetical factors (Truong, 2016). The results from the first EFA attempt identified only one factor; no rotation was possible. These results were based on the Eigenvalues of 1.0. While the results showed that a single factor accounts for 69.33% of the variance, in reviewing the scree plot, it was of interest to explore the data further through a second round of EFA analysis. The second EFA was conducted with fixed factors of three, and with the low coefficient values being suppressed below .3. This is another area of subjective decision-making in the EFA analysis process, in which each researcher could interpret the scree plot differently (Hui, 2021). The second analysis results produced two viable factors.

Ideally, EFA results would be followed up with CFA and SEM analysis to make inferences about a hypothesized model structure based on the extracted factors. Without such confirmatory analysis it would be inappropriate to make conclusions on VR technology acceptance by aircraft mechanic students. The EFA results alone do not explain the extent and direction of the relationships between the extracted factors and how those factors influence VR acceptance. However, industry practitioners could use the results of this study to better understand and implement the extracted factors within the scope of confirmatory research for technology acceptance rather than using existing theoretical models, one of which being TAM. For example, this study anticipated three distinct factors, BI, PU, and PEU, consistent with TAM. Yet, the exploration of the data revealed that there are only two factors that are plausible, of which one includes all of the BI and PU items and PEU2. These results enable researchers and industry practitioners to look beyond the theoretical framework of TAM, to generate and test research hypotheses that are based on actual data rather than theory. A data-driven approach is based on student opinions, which theoretical models lack. Also, unlike TAM and other technology independent theoretical models, the results of this study are specific to the use of VR for aircraft maintenance training. Hence, the model structure or factors/constructs extracted through EFA analysis are technology and context specific.

Research Question. Contrary to the theorized EFA model of three factors/constructs of Perceived Usefulness (PU), Perceived Ease of Use (PEU), and Behavioral Intention (BI), the final EFA results produced only two factors. Factor 1 was represented by measurement items BI1, BI2, BI3, PU1, PU2, PU3, and PEU2. Factor 2 was associated with only two measurement items, PEU1 and PEU2. Keeping in mind that the purpose of EFA is exploratory and not suitable for hypotheses testing, the results sufficiently address the research question of, "What are the underlying factors of the student aircraft mechanic survey?" with some limitations. Factually, the study was able to extract two underlying factors/constructs from the EFA analysis (through dimensional data reduction). However, some researchers suggest that each factor should have at least three measurement items in order to be considered (Hui, 2021). This is more relevant for the purpose of following up with a confirmatory analysis (CFA) to test model fit. For an exploratory study, the factor loading and correlation coefficients for Factor 2 with only two measurement items were high enough (>.7) to be considered as belonging in a separate group from the rest of the variables (see the pattern matrix and structure matrix in Figures 12 and 13). Hence, the EFA analysis results were deemed sufficient for addressing the exploratory research question.

The only challenge was to assign construct names to Factor 1 and Factor 2. Examining the groupings of the measurement items, the results suggested that participants responded in the same way to all measurements for Behavioral Intention (BI) as they did to Perceived Usefulness (PU). It is possible that these constructs are highly correlated and not differentiated from one another by the participants. In other words, if VR is perceived to be highly useful, then it is expected for the participants to respond to PU items from the perspective of having intent to use VR. Alternatively, the results could imply improvements are needed for survey instruments to better differentiate the survey items for BI and PU. More is discussed in the next section within the context of validity.

Nevertheless, the results from this study suggest that participants consistently responded in the same way to the use of VR for aircraft maintenance training as they did to VR enhancing learning, VR being useful for real world aircraft maintenance, and VR being efficient for aircraft maintenance training. From a theoretical perspective, the usefulness of VR seemed to be the main reason participants rated BI items favorably. This was evident in the comments provided, where many participants explained that their willingness to use VR is based on its usefulness in gaining skills/knowledge/experience. There was also some mention of perceived enjoyment, the *cool factor* of VR. Less favorable ratings on BI items expressed dependencies on how the technology would be implemented, or a preference for hands-on experience. For example, some participants with neutral or lower ratings stated that they would use VR "as a learning experience but not as training," they would use VR "as a supplement to experience and technical data," or as a "lecture replacement." The EFA results and observations from the qualitative data offer some practical implications for the students' adoption of VR technology for aircraft maintenance training. Since PU seems to be so closely associated with BI and represented by the same factor, the aviation industry and regulatory authorities need to consider implementation aspects of VR. Not only could a poor implementation take away from the usefulness of the technology, but it could also discourage students from using VR because it is no longer enjoyable. Similarly, how the technology is implemented, as supplemental vs. replacement of current teaching practices, could also impact the student's adoption of VR. Preserving hands-on experiences with real aircraft/aircraft parts also seems to be important to aircraft mechanic students.

The EFA analysis results in an extraction of factors and identification of the survey items that are associated with each factor. To define a model structure that could potentially be tested through full SEM analysis to confirm the relationships of the factors, and the extent and direction of those relationships, it is necessary to determine the concepts those factors represent. This is typically achieved by considering the items in the factor groups, what they have in common, and assigning a construct name based on the interpretation of the results. In this study, Factor 2 clearly included items that only represent the construct of perceived ease of use, with only two items (PEU1 and PEU3). However, Factor 1 not only included all BI and PU items but also PEU2. As previously discussed, participant responses to PEU2 suggested that the respondents considered more of the usefulness of VR rather than its ease of use. Also, it was noted that BI responses mostly considered the usefulness of VR. Therefore, Factor 2 would be representative of Perceived Usefulness as the underlying construct. This means that the EFA results provide some validation of TAM, where perceived ease of use and perceived usefulness are presented as factors that potentially contribute to technology acceptance. The study being exploratory in nature does not allow making any inferences on the acceptance of VR by students for aircraft maintenance training. However, it does provide a basis for future confirmatory research and practical considerations for how to implement the technology to make it easy to use and useful to students, such that it can facilitate adoption.

Reliability and Validity. The reliability analysis was a straightforward process using the Cronbach's Alpha method. The results showed that the measurement items for each factor extracted (Factor 1 and Factor 2) were highly correlated with each other. The reliability analysis results did not indicate a need to remove any of the measurements from EFA analysis. However, the validity of the results, as assessed through the factor correlation matrix (see Figure 14) is perhaps an area of debate. Technically, the results showed high validity, adequate distinctness of each factor, based on a factor-to-factor correlation coefficient of .687 (<.7). However, one could argue that the results are borderline adequate as the correlation value is just below .7 or exactly .7 when rounded up. Having values too close to the threshold of .7 may be indicative of a multi-collinearity issue. Updating the survey questions to prevent confusion and including more than three measurement items for each theorized construct, to be able to remove problematic items but still run the EFA analysis (retaining the preferred minimum of four measurement items), are ways to avoid this issue and improve the validity of the study (Hui, 2021). Given that this research was part of a degree program, and the scope of the study needed to be narrow, there was no opportunity to do a follow up study incorporating lessons learned.

Conclusions

The original purpose of this research was to test a proposed Technology Acceptance Model structure for acceptance of VR technology for aircraft maintenance training, by applying SEM analysis, which is confirmatiory in nature. However, due to not being able to demonstrate a good model fit with the small sample size from the pilot study, and given the small sample size for the final study, the purpose of the research was changed from confirmatory to exploratory. Both the confirmatory and exploratory analysis methods included many subjective decisons that could have influenced the results. Ultimately, the exploratory analysis was deemed to be the most appropriate approach for the final study with the sample size of 120 participants.

The results of the EFA suggested that the measurements could belong to two distinct factors/constructs. This was adequate in answering the research question but did not exactly match the theorized constructs of Behavior Intention (BI), Perceived Usefulness (PU), and Perceived Ease of Use (PEU). Additionally, the validity for factors extracted throught EFA were acceptable as suggested by the correlation coefficents being below .7 to demonstrate distinctness. Still, the factors could be considered as poorly defined with Factor 1 being assoociated with the majority of the measurement items and contributing to 66.741% of the variance, and Factor 2 including only two measurement items and accounting for a much smaller percentage of the variance (6.297%). These results suggest that the study could have been designed better to result in a more defined factor structure by including more measurement items for factor extraction, including additional constructs to be explored, and putting more effort into the design of the survey insturment with help from multiple subject matter experts (SMEs). With the actual data collected, the chances of being able to obtain a different model structure and stronger validty test results than the current results would be improbable, even if other EFA extraction and rotation methods were used. Typcially, researchers prefer to use the PCA extraction method for EFA with orthogonal rotation for practical reasons, but it is often the case that the results end up being similar to other EFA extraction methods with oblique rotation (Truong, 2016). In conclusion, the results of the final EFA analysis did produce two factors, which based on the interpreation (as covered in the previous sections) were named as Perceived Usefulness (Factor 1) and Perceived Ease of Use (Factor 2). However, there were many lessons learned that helped identify potential areas of improvements for future research and theoretical/practical implications.

Theoretical Contributions

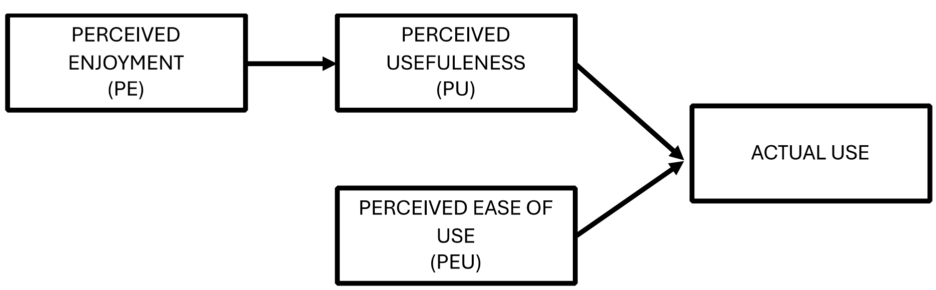
This study contributes to the existing research in many ways. First, the topic of this research is specific to aviation, and the current literature on the use of VR for aircraft maintenance training, exploring these exact constructs (PU and PEU), is nonexistent. This study contributes to the body of knowledge within the aviation industry in the area of training that could benefit from future improvements by utilizing more modern methods and tools, such as VR. The study identifies factors that could be considered in future confirmatory research and provides some insight on aircraft mechanic students' opinions. For example, the comments participants provided for rating the BI items reflected their perceived enjoyment of using VR and dependencies for VR being perceived as useful for aircraft maintenance training as associated with their intent to use

VR. From a theoretical perspective, this study suggests that VR acceptance by student aircraft mechanics could be influenced by external factors as well as technology-specific/task specific factors. The EFA results offered a model structure composed of PU and PEU constructs as potential contributing factors to actual use of VR technology for aircraft maintenance training.

Second, while the scope of this study was exploratory, the measurement items/variables were theorized based on the review of existing literature. No theoretical framework was required for exploring the data and the theorized constructs through EFA. However, given that the study retained the original constructs from the TAM model, the results could provide additional insight to other researchers that are planning to use TAM as the bases of their hypothesized model structure, specifically for constructs BI, PU, and PEU. Since the EFA results did not extract a third factor that would truly represent the construct of BI, the suggested model structure for future research is shown in Figure 17, taking into consideration a possibly plausible external variable of perceived enjoyment.

Figure 17

Proposed Model Structure Based on EFA Results and Observations of Participant Comments



Finally, the study contributes to the literature with its statistical method. There are so many different multivariate statistics methods, and EFA being one of those methods has some unique assumptions. This study contributes by serving as an example for testing EFA assumptions and conducting the analysis. The study specifically outlines the decision points and offers a way of handling data that is not normally distributed. The study demonstrates how those decision points were approached to select the most appropriate EFA extraction and rotation method. Having access to such examples in the literature is especially useful in academia.

Practical Contributions

The focus on the use of VR for aircraft maintenance training has some practical implications within the aviation industry for AMT schools, researchers, practitioners, and regulatory agencies. With student aircraft mechanic participants recruited nationwide and given the participant diversity/distribution of the participant characteristics (i.e., gender, race, education status), the study results are considered to be generalizable to the greater population of student aircraft mechanics. The results also showed themselves to be reliable and valid. Thus, the study provides insight on student aircraft mechanics' perceptions of VR technology for aircraft maintenance training. As VR gains popularity with AMT schools and is adapted for aircraft maintenance training, it will be important for institutions to consider predicted use of the technology within the context of perceived usefulness and perceived ease of use. This is because the EFA results suggested that perceived usefulness is potentially very closely associated with behavioral intention to use VR for aircraft maintenance training. Having awareness of students' opinions about VR could help educators and AMT schools implement the technology in a more effective way, by ensuring ease of use, usefulness, and carefully selecting the type of training that is most suitable for VR. For example, when rating BI items, students reported in their comments that usefulness is an influencing factor for using VR for aircraft maintenance training. Making sure that hands-on practice/experience is balanced with VR training was also an important factor in the adoption of VR. While the qualitative data was not statistically analyzed in this study, it does help the industry to understand how aircraft mechanic students think. Furthermore, the industry may want to take advantage of the reported perceived enjoyment of using VR for aircraft maintenance training by implementing VR technology in a timely manner, and perhaps phasing the implementation to not overwhelm the students. For example, AMT schools may want to introduce basic (i.e., single user) VR functionalities first for limited training tasks or lessons until students get acquainted with operating and wearing VR devices, and then introduce more advanced capabilities, such as immersive VR with multiusers for two-person tasks and even physical mockups to simulate interferences with the physical environment for maintenance tasks that are conducted in tight or hard to reach spaces or spaces that require a challenging body posture. The phased approach may help maintain the perception of VR technology for aircraft maintenance as easy to use, which the EFA results showed to be a plausible factor. Finally, the study offers many lessons for other researchers interested in the same topic, research design, and analysis methods. Having these study results documented and available to the general public may especially benefit many novice researchers in practice, by giving them an opportunity to criticize and develop ways to improve their own research with similar objectives.

Limitations of the Findings

In addition to the limitations presented in Chapter I, there were multiple specific limitations identified. Some limitations were associated with the sampling method and how the survey was administered. There was no effective way of contacting the FAA approved schools or students directly to recruit participants. The initial intent was to use judgement sampling (a nonprobabilistic method), but due to the challenge of schools that were not willing to participate in the study or share student contact information, the priority became to reach out to as many schools as possible rather than balance out between large vs. small schools, private vs. public schools, in the various geographical areas to ensure diversity. Acknowledging the challenges associated with recruiting participants, the final survey was solicited through social media as well. While there was an indication that AMT schools across the U.S. are represented to some extent,, the diversity of the sample, and in turn the generalizability of the study results to the target population could be limited. The study could not ensure that AMT programs of all sizes are representated, or even coverage geographically to ensure that the entire U.S. population of student mechanics are represented. Also, the use of social media means opening the survey up to the general public. While measures were taken in the consent form to avoid non-qualified participants, the survey highly relied on the participant's honesty. Another challenge with social media was not being able to reach out to all specific aircraft maintenance groups to target qualified participants. Often, an attempt to post to these specific groups resulted in accounts being blocked or deletion of the post. Some groups had strict criteria for joining, which made it impossible. This limited how much of a sample size could be achieved within a given time of the survery being open.

Finally, the survery may have been completed by participants that were more motivated by the incentive rather than out of interest to contribute to the study. This can limit the diversity of participant characteristics based on their motivational factors, which was not assessed.

While the small sample size presented some limitations with respect to the statisitical analysis, the results could also have been influenced by the researcher's subjectivity of the statistical techniques for decisions made every step of the way. For example, the pilot study results suggested that the sample size was too small to achieve a good model fit through CFA. However, the process to obtain a good model fit and how many times to manipulate the model and retest is based on the subjective judgement of each researcher. There is no clear way of telling when the model is considered too convoluted and where to draw the conclusions. Another area of such subjectivity was with respect to selecting a specific EFA method. With normality not being achieved, possibly due to the limited sample size, this study was limited to the PAF extraction method. It is possible that results could have been different if normality was achievable. The subjectivity is in how hard each researcher tries to get close to normality, and in the end, which EFA method is selected when there is no perfect normality. For this study, the alternative was to normalize the data as much as possible for EFA analysis with a different extraction method, to compare the results and see if they were any different. Another missed opportunity was to design the study better with more than just three measurement items for each theorized construct, and possibly include more constructs to explore. However, these decisions are so subjective and can differ from one researcher to another, that it is fair to acknowledge that the study results may have been limited by the researcher's decisions throughout the study planning and analysis process. In this case, the research decisions and reanalysis efforts were also impacted by the researcher's individual time constraints, making sure to keep the scope and retest efforts feasible for timely completion of the study.

Recommendations for Future Research

This section provides recommendations based on the study results and conclusions. The recommendations consider the target population, future research methodologies, and future areas of research. The lessons learned from this study can be very informative to other researchers who plan to conduct studies on the same topic -- acceptance of VR for aircraft maintenance training. The aviation industry would benefit from continued research in this area considering the gaps in the current literature. The topic of VR use for aircraft maintenance training is relevant not only to AMT schools but also to governing bodies, such as the FAA. The literature synthesis in this study showed that the current aircraft maintenance training requirements are outdated (White et al., 2000), and the industry is in need of more modern instructional methods and tools (Goldsby & Soulis, 2002). The most appropriate path of moving towards the adoption of VR for aircraft maintenance training in order to provide students with enhanced learning opportunities is to first have a good grasp on what the end users (student aircraft mechanics) think about VR, and how well the technology might be accepted if implemented. Thus, additional research on the same topic as this study could help industry practitioners to make more informed decisions.

Recommendations for the Target Population

This study focused on end user perceptions of VR, the end users being student aircraft mechanics. Therefore, it was important to achieve a representative sample. The publicly available data on the target population was found to be very limited, and in many cases schools refused to share enrollment information for adequate comparison of the sample demographics. In this case, the target population was students who are residents of the U.S., at least 18 years of age, who are currently enrolled in an FAA approved AMT program/school within the U.S. Based on the demographic results of this study, it is recommended for future research to consider additional demographic parameters that are clearly defined. This may allow for better representation and generalizability to the target population.

The sampling and survey solicitation methods could also be improved. This study followed a non-probabilistic strategy for sampling, which is deemed acceptable considering the large sample size requirements for SEM or EFA analysis. However, there was no way to tell if sample diversity based on the size and geographic location of the school is actually achieved with the convenience/judgement sampling approach. Some upfront research is recommended in the planning phase of future research, to identify public and private AMT schools of all sizes and from various geographic areas of the U.S. in order to ensure diversity. Furthermore, if time and budget are not an issue, it is also recommended to first create partnerships with strategically selected/sampled schools and administer the survey in-person, with support from the schools' faculty. This is one of the biggest lessons learned in the execution of this study as participation for the pilot study was easier to achieve with an in-person introduction of the survey to the students and support from ERAU faculty. For the final study, in an effort to achieve a large sample size, multiple forms of communication were used to reach out to schools and student aircraft mechanics, including social media and calling/emailing the schools and individual instructors. Yet, the outcome was not particularly great; the required sample size for confirmatory analysis was still not met. Hence, the perception from this experience is that this type of survey research, with such a specific target population, requires networking with academia (AMT schools and instructors) ahead of time.

Recommendations for Future Research Methodology

The methodology of this research was quantitative and confirmatory in nature as the initial scope was to test hypotheses based on the Technology Acceptance Model (TAM) as the theoretical framework for the constructs of Perceived Usefulness (PU), Perceived Ease of Use (PEU), and Behavioral Intention (BI). The data on students' perception of VR was collected through a survey instrument. However, the plan for conducting confirmatory analysis to test the relationships between the constructs (latent variables) was compromised due to the challenge of achieving the required sample size. Hence, the study scope and the data analysis were changed to exploratory. Based on lessons learned from this study, researchers are advised to consider other research methodologies for fulfilling the main objective of the study -- students' acceptance of VR technology for aircraft maintenance training.

One recommended approach is to run path analysis through SEM for the model structure suggested by the results of this study. This would allow researchers to test hypothesized relationships between the constructs of Perceived Enjoyment (PE), Perceived Usefulness (PU), and Perceived Ease of Use (PEU) as contributing factors to actual use of VR. To address the issue of inadequate sample size for SEM analysis, researchers could use Bootstrapping technique to resample the data. Bootstrapping would confirm that the SEM analysis results are valid. Bootstrapping creates multiple simulated datasets by randomly selecting data points from the original dataset. It is a popular technique for estimating the population distribution of means, standard deviation, etc. and is especially useful when the original dataset is skewed (Frost, 2024). It is important to note that Bootstrapping is not a guaranteed solution to skewed data and small sample size. The technique is subject to certain assumptions and limitations. Bootstrapping assumes that the original dataset/sample accurately represents the population. The technique could also render some unusable samples and less accurate results for SEM analysis and interpretation when the original sample size is under 200. Bootstrapping is also known to proliferate outliers (Kline, 2011). In this study, the sample was deemed as representative of the student aircraft mechanic population based on gender and ethnicity distributions only because no other data was available to make additional comparisons. However, if population data becomes available in the future and the current sample is confirmed to accurately represent the population, then conducting SEM with Bootstrapping would be a viable option. Finally, future confirmatory research on the exact proposed model from this study would require metric data on Perceived Enjoyment (PE) as the construct was only substantiated through participant comments rather than the survey items.

Another recommended option is qualitative research, in which students could be interviewed to provide more in-dept opinions on VR around the same constructs. The advantage of an interview method is that a smaller sample size is required for the qualitative data analysis. While interviews are more time consuming compared to a Likert scale survey method, they do present an opportunity to gather additional context and explanation for students' opinions on PEU, PU, and BI.

If the intent is to keep the survey method based on the TAM constructs, a procedural recommendation would be to conduct a hands-on VR familiarization session with the students. Of course, this would not be feasible to do with hundreds of participants but based on some of the participant comments in this study, it may be beneficial. This study showed a video prior to the survey to familiarize students with VR; however, there are indications that some participants who have never used VR before struggled with providing an opinion on PU, PEU, and BI, and selected a *neutral* rating to multiple questions. This was evident when screening their comments to the questions. These participants expressed, "I have no opinion," "Never used VR for maintenance training, so I don't know," "I have never used any VR before, so I have no comment," and so on. Nevertheless, these participants still provided ratings, perhaps just to qualify for the incentive. With this study's methodology, it would be impossible to accurately guess their intent. A hands-on familiarization session with a small aircraft maintenance task might avoid these situations by helping participants in forming and providing opinions. Otherwise, as with the current method, participants are asked to answer the questions hypothetically. Of course, the hands-on familiarization training would force the study into a smaller sample size from a feasibility perspective, which would require a different analysis approach since SEM and EFA are sensitive to the sample size. Regardless, this is one way of approaching the same problem statement as in this study but with a different methodology, which could produce more valid responses.

Recommendations for Future Areas of Research

Considering the exploratory nature of this study, there is an opportunity for other researchers to expand on the findings. The scope of this study was limited to only three theorized constructs, BI, PU, and PEU. This study did not examine other constructs that potentially contribute to acceptance of VR for aircraft maintenance training, or VR use in general. While the EFA method for data analysis does not require a foundational theory to produce a factor structure prior to exploratory analysis, the current literature on learning theories and previous research on VR/AR technology acceptance can help researchers extend the scope to many other constructs to include in the survey questionnaire. For example, researchers could include the external variables that this study failed to address. Based on the literature review conducted for this study, there is support for exploring the constructs of Perceived Behavioral Control (PBC), Perceived Enjoyment (PE), Perceived Health Risk (PHR), Performance Expectancy (PEXP), and Self-Efficacy (SE). These constructs were originally identified to be part of this study when the intent was confirmatory analysis. However, the pilot study results failed to achieve a good model fit through CFA to continue forward with including them in the final study. This does not mean that these constructs are invalid and would not contribute to the acceptance of VR technology. Instead, the pilot study results from this research suggest an opportunity for future research to first explore these constructs to establish a factor structure that can then be followed up with confirmatory analysis. Hence, the recommendation for a future study is to conduct another survey to include measurement items (more than three per construct recommended) for exploring the external factors/constructs of PBC, PE, PHR, PEXP, and SE, which this research did not address.

Then, once a factor structure is identified through EFA, the recommendation is to follow up with CFA analysis to confirm the relationships between those constructs. Future research could also expand further to conduct a full SEM analysis for understanding the direction of the confirmed relationships.

References

Al-Ahmari, A. M., Abidi, M. H., Ahmad, A., & Darmoul, S. (2016). Development of a virtual manufacturing assembly simulation system. *Advances in Mechanical Engineering*, *8*(3). https://doi.org/10.1177/1687814016639824

American Psychological Association. (2018). Heywood case. *APA dictionary of psychology*. https://dictionary.apa.org/heywood-case

Anaesth, A. C. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, *22*(1), 67-72. https://doi.org/10.4103/aca.ACA_157_18

ATEC. (2023). Pipeline report. https://assets.noviams.com/novi-file-uploads/atec/ATEC-2023-PipelineReport-20231221.pdf

Aviation Maintenance Technician Schools, 14 C.F.R. § 147 (2021). https://www.ecfr.gov/current/title-14/part-147

Aviation Maintenance Technician Schools, 87 F.R. 31391 (proposed May 24, 2022) (to be codified at 14 C.F.R. 14 § 43, 65, 147). https://www.federalregister.gov/documents/2022/05/24/2022-10367/aviation-maintenance-technician-schools

Borsci, S., Lawson, G., & Broome, S. (2015). Empirical evidence, evaluation criteria and challenges for the effectiveness of virtual and mixed reality tools for training operators of car service maintenance. *Computers in Industry*, *67*, 17-26. https://doi.org/10.1016/j.compind.2014.12.002

Boyd, D., & Stolzer, A. (2015). Causes and trends in maintenance-related accidents in FAA-certified single engine piston aircraft. *Journal of Aviation Technology and Engineering*, *5*(1), 17-24. https://doi.org/10.7771/2159-6670.1123

Brauer, M., & Klingauf, U. (2008, August 18-21). *Virtual-reality as a future training medium for civilian flight procedure training* [Paper presentation]. AIAA Modeling and Simulation Technologies Conference and Exhibit, Honolulu, Hawaii, USA. https://doi.org/10.2514/6.2008-7030

Broderick, S. (2017). *ATEC study: AMT schools need to boost enrollment.* AIN. https://www.ainonline.com/aviation-news/business-aviation/2017-12-19/atec-study-amt-schools-need-boost-enrollment

Burigat, S., & Chittaro, L. (2016). Passive and active navigation of virtual environments vs. traditional printed evacuation maps: A comparative evaluation in the aviation domain. *International Journal of Human-Computer Studies, 87*, 92-105. https://doi.org/10.1016/j.ijhcs.2015.11.004

Byrne, B. (2016). *Structural equation modeling with AMOS.* Taylor & Francis. https://doi.org/10.4324/9781315757421

CAE. (2024). *Transforming aircraft technician training with extended reality (XR)*. https://www.cae.com/media/documents/Civil_Aviation/Brochures/aircraft-maintenance-training-vr-whitepaper-2024.pdf

Chuttur, M. Y. (2009). Overview of the technology acceptance model: Origins, developments and future directions. *All Sprouts: Working Papers on Information Systems*, *9*(37). https://aisel.aisnet.org/sprouts_all/290

Clavin, W. (2019). *Virtual reality for scientists.* Caltech. https://www.caltech.edu/about/news/virtual-reality-scientists

Cooper, N., Milella, F., Cant, I., Pinto, C., White, M., & Meyer, G. (2016). Augmented cues facilitate learning transfer from virtual to real environments. In E. Veas, T. Langlotz, J. Martinez-Carranza, R. Grasset, M. Sugimoto, and A. Martín (Eds.), *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*, 194-198. https://doi.org/10.1109/ISMAR-Adjunct.2016.0075

Cromar, T. (2024). *Ogden ALC uses virtual reality to enhance aircraft maintenance training*. https://www.hill.af.mil/News/Article-Display/Article/3736890/ogden-alc-uses-virtual-reality-to-enhance-aircraft-maintenance-training

Federal Aviation Administration. (n.d.-a). *FAA-Approved aviation maintenance technician schools* [Search results for United States]. U.S. Department of Transportation. https://av-info.faa.gov/dd_sublevel.asp?Folder=%5CMechanicSchools

Federal Aviation Administration. (n.d.-b). Chapter 10: Environmental justice. *Airports Desk Reference*. U.S. Department of Transportation. https://www.faa.gov/documentLibrary/media/FAA_Desk_Reference_for_Airport_Actions_Chapter_10_EJ.pdf

Franklin Institute. (2021). *History of virtual reality*. [An online article by the Franklin Institute]

Frost, J. (2024). *Introduction to Bootstrapping in statistics with an example*. Statistics By Jim. https://statisticsbyjim.com/hypothesis-testing/bootstrapping/

Fussell, G. S. (2020). *Determinants of aviation students' intentions to use virtual reality for flight training*. [Ph.D. dissertation, Embry-Riddle Aeronautical University]. Scholarly Commons. https://commons.erau.edu/edt/542/

Goldenstein, P. (2020). Air Force turns to VR, AR for training and maintenance. *FedTech.* https://fedtechmagazine.com/article/2020/04/air-force-turns-vr-ar-training-and-maintenance

Goldsby, R. P. & Soulis, A. S. (2002). *Optimization of aviation maintenance personnel training and certification*. https://www.gao.gov/assets/gao-03-317.pdf

Gomez-Cambronero, A., Miralles, I., Tonda, A., & Remolar, I. (2023). *Immersive virtual-reality system for aircraft maintenance education: A case study*. https://doi.org/10.3390/app13085043

Grabowski, A. (2021). *Virtual reality and virtual environments: A tool for improving occupational safety and health.* CRC Press. https://doi.org/10.1201/9781003048510

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis.* Prentice Hall.

Huang, H. & Liaw, S. (2018). An analysis of learners' intentions towards virtual reality learning based on constructivist and technology acceptance approaches. *The International Review of Research in Open and Distributed Learning, 19*(1). https://doi.org/10.19173/irrodl.v19i1.2503

Hui, S. L. (2021). *Exploratory factor analysis: Part 2*. [Video]. YouTube. https://www.youtube.com/watch?v=jEdNrac8Vdo

HQSoftware. (2024). *VR in aviation maintenance training*. https://hqsoftwarelab.com/blog/virtual-reality-aircraft-maintenance-engineering

IASA. (n.d.). *Maintenance error: Major mayhem from maintenance mistakes.* [List of maintenance-related accidents and incidents]

Irvine, K. (2017). XR: VR, AR, MR – What's the difference? A definitive guide to navigating the landscape of Extended Reality. *The Vignet Newsletter*. https://www.viget.com/articles/xr-vr-ar-mr-whats-the-difference/

Jang, S., Vitale, J., Jyung, R., & Black, J. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computer & Education*, *106*, 150-165. https://doi.org/10.1016/j.compedu.2016.12.009

Jerald, J. (2016). Course: Human-centered design for VR interactions. *Proceedings of SIGGRAPH '16: ACM SIGGRAPH 2016 Courses* (pp. 1-60). . https://doi.org/10.1145/2897826.2927320

Junglas, I., Goel, L., Abraham, C., & Ives, B. (2013). The social component of information systems – How sociability contributes to technology acceptance. *Journal of the Association for Information Systems*, *14*(10), 585-616. https://doi.org/10.17705/1jais.00344

Kline, R. B. (2011). *Principles and practice of Structural Equation Modeling* (3rd ed.). The Guilford Press. http://ndl.ethernet.edu.et/bitstream/123456789/74702/1/35.pdf

Kozlak, M., Kurzeja, A., & Nawrat, A. (2013). Virtual reality technology for military and industry training programs. In A. Nawrat and Z. Kuś (Eds.), *Vision based systems for UAV applications: Studies in computational intelligence* (Vol. 481). Springer. https://doi.org/10.1007/978-3-319-00369-6_21

Lala, G. (2014). The emergence and development of the technology acceptance model (TAM). In I. Plăiaş (Ed.), *The Proceedings of the International Conference "Marketing - from Information to Decision"* (pp. 149-160). Babes Bolyai University.

Lange, K. (2020). *Virtual, augmented reality are moving warfighting forward*. U.S. Department of Defense. https://www.defense.gov/News/Feature-Stories/Story/Article/2079205/virtual-augmented-reality-are-moving-warfighting-forward/

Langley, A., Lawson, G., Hermawati, S., D'Cruz, M., Apold, J, Arit, F., & Mura, K. . (2016). Establishing the usability of a virtual training system for assembly operations within the automotive industry. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *26*(6), 667-679. https://doi.org/10.1002/hfm.20406

Levy, Y., & Green, B. D. (2009). An empirical study of computer self-efficacy and the technology acceptance model in the military: A case of a U.S. Navy combat information system. *Journal of Organizational and End User Computing*, *21*, 1-23.

Lewis, C. C., Fretwell, C. E., Ryan, J., & Parham, J. B. (2013). Faculty use of established and emerging technologies in higher education: A unified theory of acceptance and use of technology perspective. *International Journal of Higher Education*, *2*(2), 22-34. https://doi.org/10.5430/ijhe.v2n2p22

Li, S. (2023). *Design and development of aviation aircraft maintenance training platform based on VR technology*. https://doi.org/10.1016/j.procs.2023.11.118

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, *95*, 174-187. https://doi.org/10.1016/j.compedu.2016.01.001

Lowenstein, A. (2024). *Florida Tech, visionary training resources partner for classroom VR*. https://news.fit.edu/aeronautics-aviation/florida-tech-visionary-training-resources-partner-for-classroom-vr

Makowski, D., Sperduti, M., Nicolas, S., & Piolino, P. (2017). "Being there" and remembering it: Presence improves memory encoding. *Consciousness and Cognition*, *53*, 194-202. https://doi.org/10.1016/j.concog.2017.06.015

Marshall University. (2021). *Aviation maintenance technology program will be first to incorporate virtual reality painting*. https://www.mfg.marshall.edu/aviation-maintenance-technology-program-will-be-first-to-incorporate-virtual-reality-painting

Marzano, A., Friel, I., Erkoyuncu, J. A., & Court, S. (2015). Design of a virtual reality framework for maintainability and assemblability test of complex systems. *Procedia CIRP*, *37*, 242-247. https://doi.org/10.1016/j.procir.2015.08.067

Mishra, D., Akman, I., & Mishra, A. (2014). Theory of reasoned action application for green information technology acceptance. *Computers in Human Behavior, 36*, 29-40. https://doi.org/10.1016/j.chb.2014.03.030

Murugan, S. (2024). *The future of aviation training: virtual reality simulations.* https://www.linkedin.com/pulse/future-aviation-training-virtual-reality-simulations-murugan-hr6xc

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society, 12*(3), 150–162.

Pennsylvania State University. (2024). *STAT 500: 4.1.2 - Population is not normal* [Online course]. PennState Eberly College of Science. https://online.stat.psu.edu/stat500/node/511

Rankin, W. (2007). MEDA investigation process. *AERO*, *26*(Quarter 2), 15-21. http://www.exi.boeing.com/commercial/aeromagazine/articles/qtr_2_07/article_03_1.html

Raosoft. (2004). *Sample size calculator*. http://www.raosoft.com/samplesize.html

Raytheon. (2012). *Virtual reality brings Hollywood start power to public safety*

Rupasinghe, T. D., Kurz, M. E., Washburn, C., & Gramopadhye, A. K. (2011). Virtual reality training integrated curriculum: An aircraft maintenance technology (AMT) education perspective. *International Journal of Engineering Education*, *27*(4), 778-788. https://www.ijee.ie/articles/Vol27-4/11_ijee2465ns.pdf

SAE. (2016). *SAE 2016 augmented and virtual reality (AR/VR) technologies symposium* [Technical session schedule and topic description]. http://www.sae.org/calendar/techsess/247730.pdf

Sagnier, C., Loup-Escande, E., Lourdeaux, D, Thouvenin, I., & Vallery, G. (2020). User acceptance of virtual reality: An extended technology acceptance model. *International Journal of Human-Computer Interaction*, *36*(11), 993–1007. https://doi.org/10.1080/10447318.2019.1708612

Soper, D. S. (2021). *A-priori sample size calculator for structural equation models* [Software]. https://www.danielsoper.com/statcalc/calculator.aspx?id=89

Soylu, M. Y., & Lee, J. (2024). *Embracing the VR/AR era: Are college students prepared?.* Georgia Institute of Technology. https://c21u.gatech.edu/papers/embracing-vrar-era-are-college-students-prepared

Starr, L. T., Shorts, K., & Vans, M. (2024). *Interactive aviation maintenance classroom*. https://doi.org/10.2352/EI.2024.36.13.ERVR-182

Stone, R., Watts, K., & Zhong, P. (n.d.). Virtual reality integrated welder training. *Welding Journal, 90*.

SUNY - University at Buffalo. (2021). *Constructivism*. https://www.buffalo.edu/catt/teach/develop/theory/constructivism.html

Tang, Y., Yang, Y., & Shao, Y. (2019). Acceptance of online medical websites: An empirical study in China. *International Journal of Environmental Research and Public Health*, *16*(6). https://doi.org/10.3390/ijerph16060943

Truong, D. (2016). *DAV 726 Course* [Videos]. Canvas@ERAU.

Truong, D. (2018). *DAV 724 Course* [Videos]. Canvas@ERAU.

Truong, D. (2022). *Structural Equation Modeling and AMOS - Conducting Confirmatory Factor Analysis (CFA) in AMOS* [Video]. YouTube. https://youtu.be/Vc2BSt390Q0?si=HNqTp-nEFngypFrt

University of California at Los Angeles. (n.d.). *A practical introduction to factor analysis: Exploratory factor analysis.* UCLA Statistical Consulting Group. https://stats.oarc.ucla.edu/spss/seminars/introduction-to-factor-analysis/a-practical-introduction-to-factor-analysis/

Vaughan, N., Gabrys, B., & Dubey, V. N. (2016). An overview of self-adaptive technologies within virtual reality training. *Computer Science Review*, *22*, 65 – 87. https://doi.org/10.1016/j.cosrev.2016.09.001

Virtual Reality Society. (2017). *History of virtual reality.* https://www.vrs.org.uk/virtual-reality/history.html

Vogt, W. P., Gardner, C. D., Haeffele, M. L. (2012). *When to use what research design.* The Guilford Press.

Vogt, W. P., Vogt, R. E., Gardner, C. D., & Haeffele, M.L. (2014). *Selecting the right analyses for your data.* The Guilford Press.

Wang, Y., Anne A., & Ropp, T. (2016). Applying the technology acceptance model to understand aviation students' perceptions toward Augmented Reality maintenance training instruction. *International Journal of Aviation, Aeronautics, and Aerospace, 3*(4). https://doi.org/10.15394/ijaaa.2016.1144

Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications, 9*(6), 476-487. https://doi.org/10.1016/j.elerap.2010.07.003

White, C. W., Kroes, M., & Watson, J. (2000). *Aviation maintenance technician training: Training requirements for the 21st century.* Federal Aviation Administration

Wolf, E., Harrington, K., Clark, S., Miller, M. (2015). Sample size requirements for structural equation models: An evaluation of power, bias, and solution property. *Educational and Psychological Measurement, 73*(6), 913-934. https://doi.org/10.1177/0013164413495237

Appendix

Permission to Conduct Research

Embry-Riddle Aeronautical University
Application for IRB Approval
EXEMPT Determination Form

Principal Investigator: Lusine Carlsson

Other Investigators: Steven Hampton

Role: Student **Campus:** Worldwide **College:** Aviation/Aeronautics

Project Title: TECHNOLOGY ACCEPTANCE OF VIRTUAL REALITY (VR) FOR AIRCRAFT MAINTENANCE TRAINING

Review Board Use Only

Initial Reviewer: Teri Gabriel **Date:** 01/04/2022 **Approval #:** 22-067

Determination: Exempt

Dr. Beth Blickensderfer
IRB Chair Signature: Elizabeth L. Blickensderfer Digitally signed by Elizabeth L. Blickensderfer Date: 2022.01.11 15:27:39 -05'00'

Brief Description:

The purpose of this study is to use the extended Technology Acceptance Model (TAM) to investigate student aircraft mechanics' acceptance of VR technologies for their training. Participants will be asked to complete an online survey via Google Forms.

This research falls under the **EXEMPT** category as per 45 CFR 46.104:

☑ (2) Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (Applies to Subpart B [Pregnant Women, Human Fetuses and Neonates] and does not apply for Subpart C [Prisoners] except for research aimed at involving a broader subject population that only incidentally includes prisoners.)

EMBRY-RIDDLE
Aeronautical University

17 June 2023

Modification of Previously Approved IRB

Campus:	Worldwide	College:	COA
Applicant:	Lusine Carlsson	Degree Level:	Doctorate
ERAU ID:	1758636	ERAU Affiliation:	Student
Project Title:	TECHNOLOGY ACCEPTANCE OF VIRTUAL REALITY (VR) FOR AIRCRAFT MAINTENANCE TRAINING		
Principal Investigator:	Lusine Carlsson		
Submission Date:	06/17/2023		
Beginning Date:	03/17/2023		
IRB Approval #:	22-067		

Modification of Approved IRB
APPROVAL

✓ Validated to meet the criteria for Exempt or Expedited Status.

IRB Approver Signature: Teri Gabriel, IRB Director

Date of Approval: June 20, 2023

Questions

1. **Change of Protocol due to**:

Increase participation by expanding recruitment:

The reason for this modification is to increase participation by expanding recruitment methods and revising the questionnaire to remove external variables (survey questions) post pilot study analysis. External variables are those factors that were hypothesized to be associated with the original Technology Acceptance Model construct variables of Perceived Usefulness and Perceived Ease of use. The questions that are subject to removal from the study post the pilot study analysis include: Self-Efficacy, Perceived Health Risk, Perceived Enjoyment, Performance Expectancy, and Perceived Behavioral Control. These questions are being removed to scale down the scope of the study, to increase the chances of meeting required sample size.

In addition, investigators will contact FAA approved schools via phone/email to extend an invitation to participate. The following social media and forums will be contacted to recruit student aircraft mechanic participants:

- Instagram
- Twitter
- Facebook
- Reddit

The attached file includes examples of the language/content for posting. Please note, investigators may have to respond in an open format to any questions by interested parties to provide additional details about the study as requested. Investigators will also be contacting FAA approved schools or social media group moderators via direct messaging to ask for help in recruiting student participants, in which case the same survey invite will be shared or asking them to share, post/repost the survey invite with their students. To prevent unqualified individuals, the survey Consent form has been updated to clarify "...an FAA approved school within the U.S." Finally, the consent form reflects an updated shorter study duration (15 minutes, estimated) due to eliminating some questions.