COMPARISON OF SAFETY PERFORMANCES OF STUDENTS, CONSTRUCTION WORKERS AND ENGINEERS IN VIRTUAL SAFETY TRAINING: AN EVALUATION OF USING STUDENTS AS SURROGATES

by
Elif Gürbüz
B.S., Civil Engineering, Boğaziçi University, 2017

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfilment of
the requirements for the degree of
Master of Science

Graduate Program in Civil Engineering
Boğaziçi University
2020

ACKNOWLEDGEMENTS

I want to sincerely thank to Assist. Prof. Semra Çomu Yapıcı for her valuable support, assistance and attention throughout the study. I gratefully appreciate her guidance which contributes crucial values in my thesis and also broadens my horizon.

It would be very hard to come to this stage without a support of valuable people in my life. First of all, I gratefully thank to my closest friend Cansu Özerim for accompanying with me in study sessions and supporting me throughout this study. I would like to thank to my friend Gurur Gamgam for his support and patience.

I would like to thank to my family members Melda Gürbüz, Yücel Gürbüz, Ece Nesli Korkmaz and our newly joined member Nil Korkmaz who have always believed in me.

Lastly, I want to thank my colleagues who voluntarily involved in the experiments and shared their opinions and experience in health and safety procedures.

ABSTRACT

COMPARISON OF SAFETY PERFORMANCES OF STUDENTS, CONSTRUCTION WORKERS AND ENGINEERS IN VIRTUAL SAFETY TRAINING: AN EVALUATION OF USING STUDENTS AS SURROGATES

Construction safety is considered as a global concern with its high rates of fatal and non-fatal occupational accidents reported every year in the world. Construction safety trainings are proven to be one of the most powerful method to minimize these high rates. However, current safety implementations are found as insufficient and ineffective in terms of increasing the safety awareness and safety knowledge of construction workers. Therefore, more effective approaches should be adopted to enhance the safety trainings and implementations. In this regard, virtual reality has been recently applied in different kinds of safety trainings and has made significant contributions in terms of occupational safety. In the development stages of such training tools, although the target group is construction workers or engineers, students are mainly preferred as subject groups due to time, money and location considerations. This thesis aims to investigate the appropriateness of using students as surrogates for construction workers and for engineers in the development of virtual safety training tools. To be able to achieve the aim of the thesis, the newly developed V-SAFE v.2 training tool, which covers the basic scaffolding tasks coupled with relevant health and safety implementations, is adopted in the experiments. In this regard, the effectiveness of the tool on students, construction workers and engineers are also evaluated appropriately. Lastly, safety performances of the construction workers on V-SAFE v.2 are evaluated considering their diverse backgrounds.

ÖZET

SANAL ORTAM TABANLI İŞ GÜVENLİĞİ EĞİTİMİNDE ÖĞRENCİLERİN, İNŞAAT İŞÇİLERİNİN VE MÜHENDİSLERİN PERFORMANSLARININ KARŞILAŞTIRILMASI: ÖĞRENCİLERİN VEKİL OLARAK KULLANILMASI

Dünyada her yıl rapor edilen yüksek ölümlü ve ölümcül olmayan iş kazaları oranları göz önünde bulundurulduğunda inşaat güvenliği küresel bir sorun olarak kabul edilir. İnşaat güvenliği eğitimlerinin, bu yüksek oranları en aza indirmek için kullanılan en güçlü yöntemlerden biri olduğu ispatlanmıştır. Ancak mevcut inşaat güvenliği uygulamaları inşaat işçilerinin güvenlik bilincini ve bilgisini artırmak açısından yetersiz ve etkisizdir. Bu nedenle, inşaat güvenlik eğitimlerini ve uygulamalarını geliştirmek için daha etkili yaklaşımlar benimsenmelidir. Bu bağlamda sanal gerçeklik teknolojisi son zamanlarda birçok farklı eğitimde uygulanmış ve iş güvenliği açısından önemli katkılarda bulunmuştur. Bu tür eğitim araçlarının geliştirme aşamalarında, hedef grup inşaat işçisi veya mühendis olmasına rağmen, daha çok öğrenciler zaman, para ve yer hususları nedeniyle deney grupları olarak tercih edilmektedir. Bu tez, öğrencilerin inşaat işçileri ve mühendisler için sanal güvenlik eğitim araçlarının geliştirilmesinde vekil olarak kullanılmasının uygunluğunu araştırmayı amaçlamaktadır. Bu tezin amacına ulaşabilmek için, ilgili iş sağlığı ve güvenlik uygulamaları ile birlikte temel iskele kurulum aşamalarını kapsayan ve yeni geliştirilen V-SAFE v.2 eğitim aracı deneylerde kullanılmıştır. Bu bağlamda, geliştirilen bu aracın öğrenciler, inşaat işçileri ve mühendisler üzerindeki etkinliği de uygun şekilde değerlendirilmektedir. Son olarak, V-SAFE v.2'deki inşaat işçilerinin güvenlik performansları farklı geçmişleri de göz önünde bulundurularak değerlendirilmektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZET v
LIST OF FIGURES viii
LIST OF TABLES x
LIST OF SYMBOLS xii
LIST OF ACRONYMS / ABBREVIATIONS xiii
1. INTRODUCTION 1
1.1. Research Motivation 1
1.2. Background of the Research 3
1.3. Problem Definition 5
1.4. Aim and Objectives 6
1.5. Research Methodology 7
1.6. Scope and Limitations 8
2. BACKGROUND 10
2.1. Worker Safety and Contributing Factors for Accidents in Construction Sector 10
2.1.1. Main Causes of Accidents in the Construction Sector 11
2.1.2. Hazard Recognition and Risk Perception in Construction Safety 13
2.2. Safety Training Methods in the Construction Sector 14
2.2.1. Current Safety Training Practices 15
2.2.2. Evaluation of Training Methods regarding Their Effectiveness 16
2.3. New Safety Training Methods in the Construction Industry 18
2.3.1. Reality Technologies 19
2.3.2. Virtual Reality Safety Training Tools in Construction Industry 20
2.4. Appropriateness of the Use of Students as Surrogates 23
2.4.1. Use of Students as Surrogates in Managerial Studies 24
2.4.2. Use of Students as Surrogates in the Virtual Safety Training Tools 26
3. METHODOLOGY 28
3.1. Hypothesis Development 28
3.2. Introduction to V-SAFE v.2 Training Tool 30

3.2.1. Identification of Risk Sources and Development of the Visuals 31
3.2.2. Scenario Development 32
3.3. Experimental Settings 43
3.3.1. Selection of Subject Groups 44
3.3.2. Experiment Flow 45
3.3.3. Experiment Materials 46
3.4. Analysis Methods 49
3.4.1. Paired T-test 51
3.4.2. Wilcoxon Signed-Rank Test 52
3.4.3. Two-factor Mixed ANOVA 53
4. RESULTS 56
5. DISCUSSION 67
5.1. Effectiveness of V-SAFE v.2 67
5.2. Use of Students as Surrogates for Construction Workers 70
5.3. Supplementary Findings 75
5.4. Limitations and Future Directions 79
6. CONCLUSION 81
REFERENCES 85
APPENDIX A: SURVEY FORM 93
APPENDIX B: ASSUMPTION CHECKS FOR PAIRED T-TEST 94
APPENDIX C: ASSUMPTION CHECKS FOR TWO-FACTOR MIXED ANOVA 95

LIST OF FIGURES

Figure 3.1. Beginning of V-SAFE v.2 Training Module. 34

Figure 3.2. Personal Protection Equipment List in the Equipment Storage. 35

Figure 3.3. Representative Player with the Selected Personal Protection Equipment. 35

Figure 3.4. Flashing Object Pointing the Material Storage Area. 36

Figure 3.5. Warning Box to Call Workers for Help. 36

Figure 3.6. Radio Icon to Call Help. 37

Figure 3.7. Material List in the Toolbox. 37

Figure 3.8. Base Plates on the Ground. 38

Figure 3.9. Adjustable Jacks on the Base Plates. 38

Figure 3.10. H Type Scaffoldings on the Adjustable Jacks. 39

Figure 3.11. Warning Box to Lock the Pin. 39

Figure 3.12. Representative Player Standing on the Planks. 40

Figure 3.13. Parachute Seat Belt Attached to the Lifeline. 40

Figure 3.14. Selecting V-SAFE v.2 Training Module with Autonomous Agents. 43

Figure 3.15. Selecting V-SAFE v.2 Testing Module with Autonomous Agents. 43

Figure 3.16. Flow Chart of Experimental Procedure. 45

Figure 5.1. Boxplots for Pre-test and Post-test Scores of Three Subject Groups. 69

Figure 5.2. Safety Performances of Three Subject Groups on Paper-based Tests. 72

Figure 5.3. Age and V-SAFE v.2 Scores of Construction Workers. 76

Figure 5.4. Mean V-SAFE v.2 scores of Construction Workers. 77

Figure 5.5. Work Experience and V-SAFE v.2 Scores of Construction Workers. 78

Figure 5.6. Mean Ratings given by Three Subject Groups. 79

LIST OF TABLES

Table 2.1. Summary of Literature Review done by Trottier and Gordon, 2018. 25

Table 4.1. Paired Samples Statistics for Paper-based Test Scores of Students. 57

Table 4.2. Paired Samples Test for Paper-based Test Scores of Students..... 57

Table 4.3. Descriptive Statistics for Paper-based Test Scores of Construction Workers.... 58

Table 4.4. Wilcoxon Signed-Rank Test for Hypothesis 2. 58

Table 4.5. Test Statistics for Paper-based Test Scores of Construction Workers. 59

Table 4.6. Descriptive Statistics for Paper-based Test Scores of Engineers. 59

Table 4.7. Wilcoxon Signed-Rank Test for Hypothesis 3. 60

Table 4.8. Test Statistics for Paper-based Test Scores of Engineers. 60

Table 4.9. Descriptive Statistics of Paper-based Test Scores in Hypothesis 4. 62

Table 4.10. Test Results of Performance Effect in Hypothesis 4. 63

Table 4.11. Test Results of Subject Type Effect in Hypothesis 4. 63

Table 4.12. Descriptive Statistics of Paper-based Test Scores in Hypothesis 5. 65

Table 4.13. Test Results of Performance Effect in Hypothesis 5. 66

Table 4.14. Test Results of Subject Type effect in Hypothesis 5. 66

Table B.1. Test of Normality for Paper-based Test Scores of Students. 94

Table B.2. Test of Normality for Paper-based Test Scores of Construction Workers..... 94

Table B.3. Tests of Normality for Paper-based Test Scores of Engineers. 94

Table C.1. Test of Normality of Paper-based Test Scores in Hypothesis 4. 95

Table C.2. Skewness and Kurtosis Values of Paper-based Test Scores in Hypothesis 4. ... 95

Table C.3. Levene's Test of Equality in Hypothesis 4. 95

Table C.4. Test of Normality of Paper-based Test Scores in Hypothesis 5. 96

Table C.5. Skewness and Kurtosis Values of Paper-based Test Scores in Hypothesis 5. ... 96

Table C.6. Levene's Test of Equality in Hypothesis 5. 96

LIST OF SYMBOLS

df	Degree of freedom
H_0	Null hypothesis
H_1	Alternative hypothesis
Kurtosis	Value of Kurtosis
N	Number of observations
Skew	Value of Skewness
$SE_{kurtosis}$	Standard error for Kurtosis
SE_{skew}	Standard error for Skewness
t	Test statistic for the paired T-test
$Z_{kurtosis}$	Calculated z-value for Kurtosis
$Z_{skewness}$	Calculated z-value for Skewness
µ1	Population mean before observations
µ2	Population mean after observations
µd	Population mean of difference

LIST OF ACRONYMS / ABBREVIATIONS

3D	Three-Dimensional
4D CAD	Four-Dimensional Computer-Aided Design
ANOVA	Analysis of Variance
AR	Augmented Reality
BAP	Boğaziçi University Scientific Research Project
BLS	Bureau of Labor Statistics
CAD	Computer-Aided Design
CFOI	Census of Fatal Occupational Injuries
df	Degree of Freedom
EU	European Union
Eur-Lex	European Legislation Database
Eurofound	European Foundation for the Improvement of Living and Working Conditions
Eurostat	Statistical Authority for the European Union
HSE	Health, Safety and Environment
ILO	International Labor Organization
IVE	Immersive Virtual Environment
MBA	Master of Business Administration
MR	Mixed Reality
NPC	Non-Playable Character
OSHA	Occupational Safety and Health Administration
PPE	Personal Protection Equipment
RE	Reality Environments
SE	Standard Error
Sig.	Significance Level
SPSS	Statistical Package for the Social Sciences
SGK	Social Security Institution
TUBITAK	Scientific and Technological Research Council of Turkey
UK	United Kingdom

VE	Virtual Environments
VR	Virtual Reality
V-SAFE v.2	Virtual Safety Training Tool version 2

1. INTRODUCTION

This section consists of six subsections by presenting the motivation of research and relevant studies in the literature, stating the problem, establishing the aim and objectives of the research, summarizing the research methodology and lastly stating the scope and limitations.

1.1. Research Motivation

Construction sector is a complex and dynamic sector involving many risk factors and therefore facing a high number of accidents and fatalities worldwide. Construction sector is one of the sectors in which occupational risks are most frequently occurred (Eurofound and ILO, 2019). The statistical authority for the European Union (Eurostat) reported that, the construction sector is the leading sector in terms of fatal accidents at work (Eurostat, 2018). According to the number of accidents occurred in construction industry for the developed countries, the construction workers face fatalities three or four times more than other workers at work and this ratio becomes three to six times greater for the developing countries (Eurofound and ILO, 2019). Construction sector is the leading industry in terms of the occupational accidents and fatalities in Turkey as well. According to statistics book in 2017 of Turkish Social Security Institution (SGK), construction sector constitutes of 17.5% of all reported occupational activities and 36% of all fatal activities in all industries in Turkey (SGK, 2017). With high rates of occupational accidents and fatalities in the construction sector, construction safety is a global concern (Guo *et al.*, 2012). Many accidents can be prevented by certain procedures and regulations facilitating the construction safety (Helander, 1991). Construction companies adopt to safety regulations based on the safety trainings required by laws (Xie *et al.*, 2006). Implementation of safety regulations become priority for construction companies (Wilkins, 2011).

In Turkey, four main laws regulate working conditions: The Labor Act of Turkey (Law No. 4857), Act No.6356 on Trade Unions and Collective Labor Agreements, Act No. 4688 on Public Servants' Trade Unions, and Act No. 6331 on Occupational Health and Safety (Eurofound and ILO, 2019). Act No. 6331 on Occupational Health and Safety is the

key guide for the health and the safety of workers in almost all sectors. In accordance with Turkish Labor Act No. 6331, workers are not allowed to work in heavy and dangerous tasks unless they get corresponding training certificates. Turkish Labor Act No. 6331 and Framework Directive are consistent with each other in terms of their training principles (Karaca and Gökçek, 2015). Turkish Labor Act no. 6331 is applied to all workplaces, in both public and private sector, for employers of these workplaces and all workers including interns regardless of their field of activity (Karaca and Gökçek, 2015). However, regulations are not implemented properly because contractors neglect the basic safety rules which leads to occupational accidents (Gürcanlı and Müngen, 2013). Gürcanlı and Müngen (2013) stated that this ignorance stems from the lack of safety culture and inadequate governmental inspections.

Health and safety training methods may differ in companies, however, all of them should ensure the conditions in correspondent regulations. The regulations on this issue include the requirements at the minimum level, therefore the companies should evaluate their health and safety risks in detail and provide the quality training to their employees (Karaca and Gökçek, 2015). Health and Safety professionals in Turkish construction companies are responsible for preparing a Health and Safety Framework which is compatible with the Environmental and Social Management Plan, and Turkish Occupational Health and Safety regulations. According to the Turkish Regulation on Work Health and Safety, an employer shall provide workers with the correspondent health and safety trainings about the work activities, workplaces, potential risks, protective and preventive measures (Turkish Official Gazette no. 28648 dated 15.05.2013). Companies provide workers with safety trainings in various ways such as readings, safety training lectures, videotapes or online classes (Xie *et al.*, 2006). Considering the high rates of fatalities and accidents at construction sites, the effectiveness of current safety training tools has been questioned.

The following sections covers the similar studies in the literature and provides new approaches for the development of health and safety trainings and implementations in the construction sector.

1.2. Background of the Research

In terms of human safety, construction industry is one of the most dangerous industries (Irizarry and Abraham, 2005) with the highest rate of accidents (Sawacha *et al.*, 1999). Although improvements in safety regulations have been made recently, safety problems are still high in the construction industry. Therefore, many researches focused on the main reasons of the high rates of accidents and fatalities in the worldwide construction sector (e.g. Abdelhamid and Everett, 2000; Haslam *et al.*, 2005; Choudhry *et al.*, 2008). Worker training, worker attitude and management procedures are three main issues to be considered as corrective actions to minimize the accidents in the construction sites (Abdelhamid and Everett, 2000). Haslam *et al.* (2005) also found that problems related to workers, specifically worker actions and worker behavior constitute of two thirds of the accidents. Ineffective hazard identification creates unsafe site conditions and construction process for workers in the construction site (Li *et al.*, 2012). Unsafe behavior of workers mainly stems from poor hazard identification and safety risk perception (Tixier *et al.*, 2014; Perlman *et al.*, 2014). Hazard recognition and risk perception skills can be effectively enhanced through safety trainings (e.g. Perlman *et al.*, 2014; Namian *et al.* 2016).

Although companies apply safety prevention measures as corrective actions, workers still perform unsafe acts (Sacks *et al.*, 2013). Lack of proper training is a leading factor for unsafe work behavior (Li *et al.*, 2012). Non-compliance with these procedures coupled with inadequate and unengaging training are the key reasons on this issue (Wilkins, 2011). Worker training and professional development are one of the key factors in enhancing the safety performance of workers in the construction sites (Li *et al.*, 2012). Despite alternative ways of safety training, workers still suffer from lack of essential safety knowledge which shows the deficiency in current safety trainings (Haslam *et al.*, 2005). Most construction safety programs adopt classroom lectures which do not fully engage workers (Wilkins, 2011). In this regard, more effective methods should be adopted to engage trainees instead of traditional classroom trainings (Wilkins, 2011).

Some researchers have focused on how to improve the quality of safety training for workers instead of the traditional methods (e.g. Li *et al.*, 2012; Guo *et al.*, 2012; Sacks *et al.*, 2013; Bhoir and Esmaeili, 2015; Namian *et al.*, 2016; Mohd *et al.*, 2019). According to the

survey among construction workers done by Wilkins (2011), there was a considerable dissatisfaction in the courses that were given to the workers. Current safety training methods include theoretical rather than being more practical-based practices and hands-on or practical based approach should be adopted in in safety trainings (Mohd *et al.*, 2019). According to the meta-analysis study by Burke *et al.* (2006), training methods are categorized into three groups based on the learners' participation in trainings: (i) least engaging (lectures and videos); (ii) moderately engaging (programmed instruction and feedback interventions), and (iii) high engaging (training in behavioral modelling and hands-on demonstrations). They found that high-engagement training was approximately three times as effective as low-engagement training (Burke *et al.*, 2006). Computer technologies increase workers' involvement and interest towards health and safety issues (Becker and Morawetz, 2004). Consistent with this finding, Sacks *et al.* (2013) found the superiority of virtual reality safety training tools over the conventional methods in terms of construction workers' learning on construction safety risks. In other words, the effectiveness of the training method is found as positively correlated with the level of engagement (Sacks *et al.*, 2013). Namian *et al.* (2016) compared high engaging and low engaging training methods in terms of their association with the level of hazard recognition and risk perception. Results showed that high-engaging trainings provide higher levels of hazard recognition and risk perception compared to low-engaging training methods (Namian *et al.*, 2016). In addition, Perlman *et al.* (2014) showed that trainees can identify hazards better in a virtual environment comparing with the traditional documents and visuals.

Among the new technologies, virtual reality environments are becoming more widely used for educational purposes (Bhoir and Esmaeili, 2015). Virtual reality technology has been successfully applied in many different disciplines such as education, health and engineering design studies and has made significant contributions to these sectors in terms of occupational safety (Lin *et al.*, 2011). Using Virtual Reality (VR) to enhance the safety training enables trainees to experience in a safe environment (Xie *et al.*, 2006). Some studies adopted virtual environments in the development of safety tools such as the 3D interactive safety assessment (Li *et al.*, 2012), a virtual construction environment for steel erection (Irizarry and Abraham, 2005), a multi-user virtual reality system with motion tracking function (Shi *et al.*, 2018) and a VR-based safety training system for construction workers' electrical hazards awareness enhancement (Zhao and Ye, 2012).

In the experiments performed with virtual reality safety training tools, students are generally used as subject group (e.g. Becker and Morawetz, 2004; Irizarry and Abraham, 2005; Xie *et al.*, 2006; Perlman *et al.*, 2014; Shi *et al.*, 2018). Becker and Morawetz (2004) conducted a survey with students to evaluate the worker activities before and after the training regarding the use of chemical spills. Irizarry and Abraham (2005) conducted two experiments on the VR safety application in the steel erection process in which a student subject group was used for the first experiment and ironworker subject group was used for the second experiment. Perlman *et al.* (2014) conducted experiments with students and professionals from the construction industry as surrogates for the construction workers. They compared the performances of each group and noted that the direct evaluation of workers' perception would be important for the future (Perlman *et al.*, 2014).

Researchers generally use student subject groups in the development of new safety training tools; however, limited research has questioned whether the students are good surrogates for the construction workers. Use of students as surrogates is considered as quick and suitable sources for researchers in terms of time, distance and money issues (Shuptrine, 1975). However, the appropriateness of using students as surrogates has been questioning and the results are inconclusive (Trottier and Gordon, 2018). Most of the studies in behavioral accounting and management information systems use students as surrogates as well. (Ashton and Kramer ,1980; Hughes and Gibson, 1991). For example, Abdel-Khalik (1974) used 40 MBA students in the experiment to compare with the managers and concluded that students are not appropriate surrogates for the managers. On the other hand, Ashton and Kramer (1980) conducted a study with students and auditors and found that students are adequate surrogates for auditors. Therefore, more research on this subject should be performed for comprehensive result (Trottier and Gordon, 2018).

1.3. Problem Definition

Most of the construction safety trainings include classroom lectures which do not fully engage the construction workers (Wilkins, 2011). More effective approaches should be adopted to enhance the safety trainings for construction workers. In this regard, virtual environments are widely used for construction safety trainings (Kassem *et al.*, 2017). Newly developed safety training tools are used not only for the construction workers but also for

engineers and intern students at the construction site. In this regard, the effectiveness of such tools on participants' learning should be tested for these three groups however there is a limited evidence on the effectiveness of virtual environment-based safety training tools on these groups.

The establishment of a new safety training tool requires a considerable time due to many adjustments and revisions occurring in the development stage of the tool. Due to the long back and forth processes, the experiments are conducted all over again for many times. Although the target group for these trainings are the construction workers or engineers, it is challenging to schedule the experiments with them due to time, location, and scheduling considerations. Therefore, before scheduling the experiments with the construction workers and engineers, developed safety training tools should be an end-product and already tested with an appropriate experiment group. Experiment groups mainly constitute of students because it is considered as a quick and convenient source of information in terms of time, distance and money (Shuptrine, 1975). There are some managerial studies proposing the appropriateness of students as surrogates (e.g. Abdel-Khalik, 1974; Ashton and Kramer, 1980; Hugher and Gibson,1991). Although there are also some researches on safety training tools adopting students as a subject group as well, there is still a gap in the literature about the appropriateness of using students as surrogates for the construction workers and engineers in the development of safety training tools.

1.4. Aim and Objectives

The use of students as surrogates is commonly adopted in the development of construction safety training tools, however, there is a gap in the literature regarding the appropriateness of this phenomenon. To fill this gap, this thesis aims to investigate the appropriateness of using students as surrogates for construction workers and for engineers in the development of virtual environment-based safety training tools.

To fulfil the main aim of this thesis, the objectives are as the following:

- Comparing the paper-based test results of each group before and after receiving the virtual safety training tool called V-SAFE v.2 in order to evaluate the effectiveness of the tool on subject groups' learning.
- Analyzing the safety performances of students with construction workers by using their paper-based test results in order to investigate the appropriateness of using students as surrogates for construction workers.
- Analyzing the safety performances of students with engineers by using their paper-based test results in order to investigate the appropriateness of using students as surrogates for engineers.

1.5. Research Methodology

This thesis investigates the following research questions: (i) Are the safety trainings developed in virtual environments effective on enhancing safety-related knowledge and performances of students, construction workers and engineers? (ii) Are students appropriate surrogates for construction workers and for engineers in the development of safety training tools?

To address these research questions, five hypotheses are developed. collecting methodologies used in the previous studies. First three hypotheses investigate the effectiveness of virtual safety training tool on the learning of three subject groups: (i) students, (ii) construction workers and (iii) engineers, separately. Other two hypotheses investigate the appropriateness of using students as surrogates for construction workers in the development of virtual safety training tools. The experiments are designed to test these hypotheses appropriately. Three subject groups are involved in the experiments: (i) 20 undergraduate civil engineering students; (ii) 20 construction workers; and (iii) 20 engineers.

The experiment materials consist of: (i) mini questionnaire including a consent form; (ii) a paper-based test covering basic information of scaffolding installation tasks and relative health and safety issues; (iii) a virtual safety training tool called V-SAFE v.2

including training and testing modules and, (iv) a survey based on the Kirkpatrick model. The experimental flow is determined as the following:

- Gathering information about participants' background via mini questionnaire.
- Measuring the initial knowledge of participants on basic scaffolding tasks and relative health and safety implementations via paper-based test.
- Providing participants with training module of virtual environment-based safety training tool on basic scaffolding tasks and health and safety implementations.
- Providing participants with testing module of the tool to record their safety performances.
- Measuring the knowledge of participants on corresponding information given in the tool via paper-based tests.
- Performing survey to observe feedbacks from participants regarding the tool.
- Analyzing the data gathered during the experiments and interpreting meaningful results by appropriate statistical analysis methods.

The data collected through V-SAFE v.2, paper-based tests, mini questionnaire and surveys are stored in Excel documents for each participant. The statistical analysis for each hypothesis is determined by considering the characteristics of the data and previous studies in the literature. As a result, paired t-test, Wilcoxon signed-rank test and two-factor mixed ANOVA are adopted to test the null hypotheses. These statistical analyses are implemented in SPSS statistical package and the interpretations are made based on the results.

1.6. Scope and Limitations

This thesis investigates the appropriateness of using students as surrogates for construction workers and for engineers in the development of virtual safety training tools. While doing this, the effectiveness of virtual safety training tool is evaluated on three subject groups' learnings: (i) 20 construction workers; (ii) 20 civil engineering students; and (iii) 20 engineers.

This thesis has several limitations which might be addressed in future research. First limitation is the selection of 20 construction workers and 20 engineers from two construction

projects. The number of construction workers and engineers is enough to test the hypotheses. However, it would be more representative data to select the subject group from more diverse construction projects.

The second limitation is related to visuals used in the virtual safety training tool due to economic considerations. The visuals do not fully represent the dynamic and complex nature of the construction environment. However, since the scope of the evaluation in the participants' learning covers only the basic scaffolding installation and related health and safety issues rather than all risks at the construction sites, the tool enables us to assess their relative performance on this scope.

2. BACKGROUND

Relevant studies regarding the construction safety trainings and the use of students as surrogates have been comprehensively stated in this section. Literature review is made under four subsections. The importance of worker safety and root causes for occupational accidents and fatalities in construction sector are emphasized by providing the relevant studies in the first subsection. Safety training methods including the current and new approaches are explained in the second and third subsections respectively. In the last section, studies adopting the use of students as surrogates phenomenon are stated accordingly.

2.1. Worker Safety and Contributing Factors for Accidents in Construction Sector

Construction sector plays an important role with its contribution to economic growth and employment of any nation. According to the statistics provided by Social Security Institution (SGK) of Turkey in 2017, the construction companies accounts for 11.4% of the overall companies and the construction workers comprise of 14.4% of the all workers among other industries in Turkey (SGK, 2017). In 2002, there were 96.578 construction companies and 713.629 construction workers in Turkey while the number of construction companies increased to 214.144 and the number of construction workers reached to 2.083.438 in 2017 (SGK, 2017). With the demand in more infrastructure and facilities resulting from the increased level of development and urbanization, the construction industry tends to expand worldwide (Awwad *et al.*, 2016).

While construction industry is rapidly expanding over recent years, the safety issues in this industry needs to be paid more attention in a broader sense. Although technological developments in health and safety issues have been made in the construction sector, this sector is still considered as one of the most dangerous sectors (Irizarry and Abraham, 2005). It involves many risks and dangers which generally end up with major accidents and fatalities. Research on construction accidents revealed that construction sector has the highest rate of occupational accidents around the world (Sawacha *et al.*, 1999). According to the number of accidents occurred in the construction sites in developed countries,

construction workers have a risk to die resulting from accident at work three or four times more likely than other workers this ratio becomes three to six times greater for the developing countries (Eurofound and ILO, 2019).

In the United States, the Bureau of Labor Statistics (BLS, 2017) Census of Fatal Occupational Injuries (CFOI) reported 971 fatal work accidents per 100,000 full-time workers at the construction site in 2017 which is the highest number among other sectors. When it comes to non-fatal occupational injury rates in private industry, the construction industry is ranked as 6th place having 194.3 thousand of non-fatal occupational injuries reported in 2017 (BLS, 2018). The statistical authority for the European Union (Eurostat) reported that, the construction sector is the leading industry in terms of fatal accidents at work (Eurostat, 2018).

Construction sector is at the top of the list in terms of its great number of occupational accidents in Turkey as well. According to SGK Statistics in 2017, 62,802 occupational accidents occurred in the construction site are reported in Turkey in 2017 which constitutes of 17.5% of all reported occupational accidents. In total of 1.633 reported fatal accidents, 587 of them occurred in construction sites which corresponds to 36% of the total fatal accidents in all industries (SGK, 2017). However, these numbers might be even higher since data collection is not quite appropriate in Turkey due to considerably high number of unregistered workers (Gürcanlı and Müngen, 2013). Considering the high accident and fatality rates in construction industry worldwide, many researchers have focused on the root causes to tackle with this global concern (e.g. Abdelhamid and Everett, 2000; Haslam *et al.*, 2005; Bhoir and Esmaeili, 2015).

2.1.1. Main Causes of Accidents in the Construction Sector

Construction sector involves many hazardous working methods and machines for workers in workplace (Helander, 1991). Therefore, there are numerous types of accidents occurring at construction sites; falls from height, electrical hazards, contact with chemicals, fire and so on. Falls are the leading cause of serious injuries and fatalities among all accidents in the construction site (Hu *et al.*, 2011). In 2017, falls from height accounted for 48.0% of the occupational fatalities in the United Kingdom, 17.0% in the United States, and 44.4% in

Turkey (BLS, 2017; HSE, 2017; SGK 2017). As stated in occupational health and safety in scaffolding studies prepared by Ministry of Family, Labor and Social Security in Turkey, falls from height is the leading accident type and following accident types such as electric shock, material drop, building machine accidents, collapse of the building part and traffic accidents on the construction site are the other common types of accidents in Turkish construction sector. Similarly, construction statistics in Great Britain done by HSE in 2019 showed that apart from falls from height, trapped by something collapsing, struck by object and struck by moving vehicle are the most common ones (HSE, 2019).

To be able to minimize occupational accidents in the construction sector, researchers have focused on the main contributors to the occupational accidents. In this regard, Abdelhamid and Everett (2000) classified the root causes of the occupational accidents into three groups: (1) failing to determine an unsafe condition; (2) proceeding an unsafe activity after recognizing an unsafe condition; (3) performing unsafe acts regardless of initial conditions. More precisely, ‘‘unsafe conditions’’, ‘‘worker response to unsafe conditions’’, and ‘‘worker unsafe acts’’ are the core elements for this model. Workers actions, attitudes, capabilities or any kind of problem from workers constitute of two thirds of the accidents (Haslam *et al.*, 2005). Recent root cause of unsafe behavior is social influence created at the construction site which is so-called safety culture (Shi *et al*, 2018). However, safety climate exists as a result of safety culture (Choudhry *et al.*, 2007). Choudhry *et al.* (2007) made a literature review on safety culture and found that safety culture would reflect the safety management system of any construction site. Meaning that, a positive safety culture can be seen as a useful tool in improving the safety issues in construction sites (Choudhry *et al.*, 2007). One of the major reasons behind the construction accidents is workers’ failure to identify hazardous situations (Bhoir and Esmaeili, 2015). Haslam *et al.* (2005) also demonstrated the reasons of workers' unsafe acts:

- safety issues have been missed in the scope of heavy workloads
- workers have a tendency to take shortcuts
- misperception of safety risk, a feeling of " it won't happen to me"

Choudhry *et al.* (2008) studied on why construction workers perform unsafe behavior by investigating 11 factors that have an impact on workers' safety behavior at construction

sites. Their study revealed that management involvement is the most effective factor for enhancing site safety (Choudhry *et al.*, 2008). Several studies show that management's support and attitudes towards safety has a positive impact on a satisfactory safety level (e.g. Sawacha *et al.*, 1999; Haslam *et al.*, 2005). For example, Haslam *et al.* (2005) demonstrated from the interviews that construction supervisors do not have comprehensive safety awareness and a broad sense of accident prevention. Site inductions and tool-box talks are common examples where the trainings are delivered by trainers with a poor understanding of safety awareness (Haslam *et al.*, 2005). Accidents stemmed from human errors mainly result from carelessness or lack of awareness of the construction workers (Abdelhamid and Everett, 2000). In this regard, Sawacha *et al.* (1999) found that superintendent's attitude towards safety directly affects workers' behavior on site and summarized the root causes behind the unsafe behavior derived as a result of the interviews with the workers as: ignorance of safety knowledge, failure to follow safety procedures, work pressure, psychological factors, and lack of safety training.

2.1.2. Hazard Recognition and Risk Perception in Construction Safety

Many studies have been carried out to identify the above-mentioned factors leading to occupational accidents in construction sector. Some causes are related to unsafe behavior of workers which stems from poor hazard identification and safety risk perception (Tixier *et al.*, 2014; Perlman *et al.*, 2014). Recent research showed that most construction hazards remain unrecognized in construction projects (Namian *et al.*, 2016). Ineffective hazard identification in turn creates unsafe site conditions and construction process for workers in the construction site (Li *et al.*, 2012). The term hazard identification is used as the term hazard recognition in some research papers (e.g. Namian *et al.*, 2016; Helander, 1991). Hazard identification and assessment are fundamental for any construction safety program (Perlman *et al.*, 2014; Namian *et al.*, 2016). The reason is that, if hazards cannot be recognized or safety risk cannot be perceived, workers might not be able to follow safety procedures effectively (Albert *et al.*, 2014). Carter and Smith (2006) pointed out the importance of hazard identification in construction safety management by investigating the safety hazard identification within the U.K. construction industry. They proposed that unidentified hazards hinder the risk assessment process and control measures to be developed (Carter and Smith, 2006). Many researchers included the hazard identification

process into their studies (e.g. Lin *et al.*, 2011; Chen *et al.*, 2013; Bhoir and Esmaeili, 2015). Lin *et al.* (2011) for example, developed a safety inspector to identify hazards embedded in a virtual environment and Chen *et al.* (2013) also included hazard identification in the augmented virtual environment safety program. To enhance the ability to capture hazardous situations beforehand via training programs is an essential step for proactive safety management (Bhoir and Esmaeili, 2015).

Apart from the importance of hazard identification on construction safety, safety risk perception is also quite important for maintaining effective safety practices (Namian *et al.*, 2016). One of the root causes derived by Abdelhamid and Everett (2000) is that proceeding an unsafe activity after identifying an existing unsafe condition which is generally performed by risk-taking workers. Construction workers generally underestimate the existing risks and hazards (Helander, 1991). When workers do not realize the importance of safety risk and reject to adopting it, they become more likely to perform risk-taking behavior (Tixier *et al.*, 2014). For instance, Choudhry *et al.* (2008) observed risk taking behavior of workers at the construction site while the workers performed risky acts to be seen as a 'tough guy' or to get a high position in the organization. Shin *et al.* (2014) proposed that there is a positive correlation between the injury experience and safety risk perception of workers. Namian *et al.* (2016) concluded that the likelihood of occupational accidents dramatically increases in cases where hazards are not properly identified, or safety risk perception is underestimated. As it is mentioned in the previous part, attitude of superintendents has a direct impact on the workers' behavior towards the safety issues (Perlman *et al.*, 2014). Therefore, Perlman *et al.* (2014) proposed that once construction superintendents' hazard recognition and risk perception skills are improved, the overall safety at the construction site should be improved as well. In this regard, many studies have also shown that these skills can be effectively improved through trainings (e.g. Perlman *et al.*, 2014 and Namian *et al.*, 2016).

2.2. Safety Training Methods in the Construction Sector

Worker trainings and professional development are deemed as a core element in reducing the number of occupational accidents (Li *et al.*, 2012). Insufficient safety trainings for the construction workers considerably contribute to the occupational accidents in the construction sector (Abdelhamid and Everett, 2000). Education and worker training are two

of main approaches for a safe work environment (Kassem *et al.*, 2017). It is crucial to find the motivational issues in safety trainings to change the attitude of construction workers (Helander, 1991). Occupational accidents in the construction site take place because of either due to lack of knowledge, training or supervision, lack of means of conducting the task safely or due to carelessness (Sawacha *et al.*, 1999). In terms of accidents associated with machinery, their study revealed that operatives believe that lack of training and skill on using machinery such as ladders and scaffolding are the main reason of injuries (Sawacha *et al.*, 1999). Wilkins (2011) also demonstrated the dissatisfaction of the current training tools and the need of more understandable safety training for workers by conducting a survey with 105 construction personnel who took the OSHA 10-hour "Construction Safety Training Course".

Construction safety is improved by increasing workers' awareness on safety issues and therefore safety training is one of the best approaches to the prevention of accidents in construction sector (Guo *et al*, 2012; Sacks *et al.*, 2013). Even though companies pay attention to precautions and safety implementations, workers can still perform unsafe acts (Sacks *et al.*, 2013). Moreover, poor hazard recognition and safety risk perception of workers leads to poor construction safety at the construction site. Workers' ability to perceive risks and to recognize the hazards and its results are determinant factors of their behavior (Sacks *et al.*, 2013; Shi *et al.*, 2018). To improve hazard recognition and safety risk perception, companies should adopt effective safety training methods which are well structured for workers (Namian *et al*, 2016). Haslam *et al.* (2005) demonstrated the contributing factors of unsafe acts in construction site and found that inadequate safety knowledge coupled with deficiencies in education and training are the main underlying reasons for workers' unsafe acts. Since safety training can be developed more content specific and therefore can be adopted as a guide for workers to how safe act should be performed (Haslam *et al.*, 2005).

2.2.1. Current Safety Training Practices

Construction workers are mainly obliged to take some of these four training options: classroom lectures, hands-on instruction, testing and other off-site trainings (Li *et al.*, 2015). One day long or even shorter of safety training is considered as sufficient to get the relevant certification for construction workers in the UK, Australia and Hong Kong (Li *et al.*, 2012).

This reveals that construction safety training requirements are still not sufficient (Sacks *et al.*, 2013). Nowadays, classroom lectures are mainly prepared as presentation including health and safety procedures and some of them are specialized for certain type of subjects such as safety on dumper truck operations (Li *et al.*, 2015). Safety trainings are provided by the companies in various ways. The most common methods that the companies adopt are as follows (Xie *et al.*, 2006):

- Reading the health and safety booklet provided by the institutions;
- Attending to safety training lectures held by the health and safety supervisors;
- Watching videos, such as "Toolbox Safety Talks"
- Attending online courses, for example the online OSHA (Occupational Safety and Health Administration) class in the United States.

Safety training is mainly delivered by trainers with a poor understanding of learning and it generally happens in the site inductions and toolbox talks (Haslam *et al.*, 2005). Guo *et al.* (2012) summarized the reason of still having high accident rate in the industry that: (i) the construction site includes many complexities in terms of equipment, materials and plants and (ii) the ineffective safety training tools having hands-off off-site environment in which the workers can not involve and only listen and watch.

2.2.2. Evaluation of Training Methods regarding Their Effectiveness

Safety training methods has been questioning in terms of their effectiveness on workers' safety awareness. For example, Wilkins (2011) has questioned the effectiveness of current construction safety training methods in the United States by conducting a survey of 105 construction personnel taken the OSHA 10-hour 'Construction Safety Training Course'. He revealed that there was a considerable dissatisfaction in the courses that were given to the workers (Wilkins, 2011). Many research areas have focused on how to improve the quality of safety training for workers instead of the traditional methods (e.g. Li *et al.*, 2012; Guo *et al.*, 2012; Sacks *et al.*, 2013; Bhoir and Esmaeili, 2015; Namian *et al.*, 2016; Mohd *et al.*, 2019). Dale (1969) classified the teaching methods used in the educational process by summarizing them in a cone named in the literature as "Dale's Cone of Experience" (Figure 2.1). Dale's cone of experience is designed as a visual model to summarize different types

of mediated learning experiences varying from most concrete experiences (at the bottom of the cone) to most abstract experiences (at the top of the cone) (Dale, 1969). According to Dale's cone of experience, the most effective way of learning is through doing the real thing or simulating a real experience (Dale, 1969) as shown in Figure 2.1.

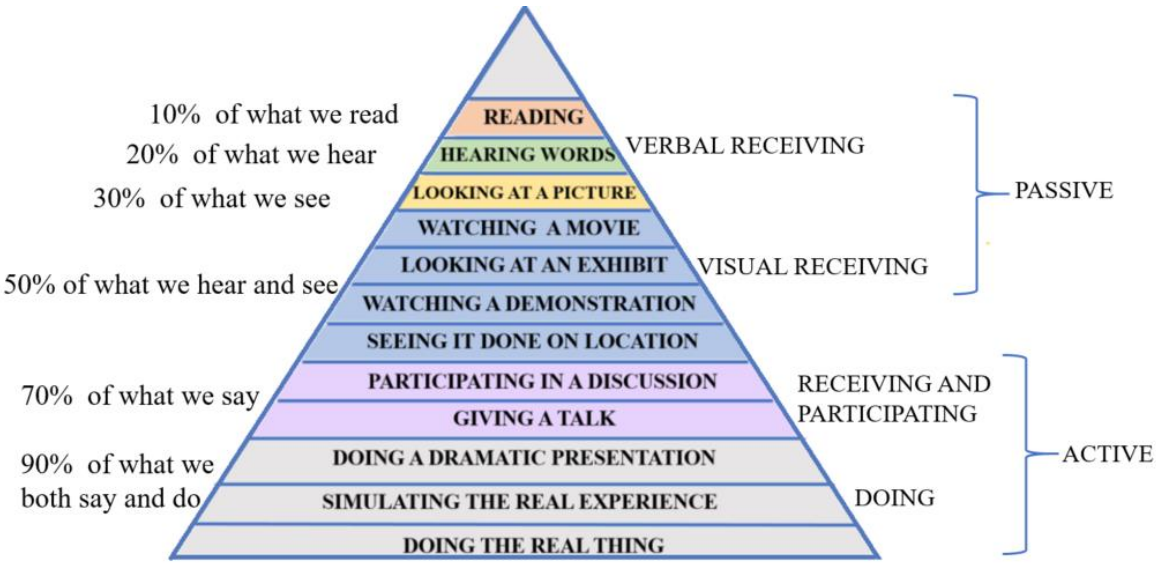

Figure 2.1. Dale's cone of experience (Dale, 1969).

Namian *et al.* (2016) proposed that unengaging training methods have hindered the effectiveness of learning and compared the high engaging and low engaging training methods on the effect of workers' hazard recognition and risk perception. Their study revealed that high engaging training tools provides with higher levels of hazard recognition and risk perception. Perlman *et al.* (2014) also drew a conclusion from their experiment that trainees can identify hazards better in a virtual environment comparing with the traditional documents and photographs.

Li *et al.* (2012) used a virtual safety training system for tower crane dismantlement which provides its participants with practicing the dismantling procedure and it was found that the performance of the participants in the proposed system was higher than the participants trained by the traditional method. Mohd *et al.* (2019) defend that current safety training methods are more theoretical rather than being more practical-based and pointed out the importance of hands-on or practical based approach in safety trainings. Since the

practical skills can be developed through on-site experience, on-site hands-on practice should be provided for the construction workers to deal with the construction risks that they face (Guo *et al.*, 2012). Practice is essential in effective training however job-site training is dangerous in construction site, since there is a high risk of injury in case of any mistake during learning and testing (Zhao and Ye, 2012). Additionally, Li *et al.* (2015) drew attention to the inefficiency of on-site training which interferes with progress on site, while off-site training does not effectively develop the workers' practical skills and awareness needed through hands-on experience.

Burke *et al.* (2006) have examined the effectiveness of various kind of health and safety training methods by conducting a meta-analysis study. They focused on three types of intervention methods which have been categorized based on the learners' engagement in the training procedure: least engaging which covers information-based trainings (lecture and videos), moderately engaging (programmed instruction, feedback mechanisms), and most engaging which focus on behavior modeling and require active participation from the learners (training in behavioral modeling, hands-on demonstrations). They found that high-engagement training was approximately three times as effective as low-engagement training. Similar meta-analysis study has been also repeated for the years 1996–2005 and this study have verified the former result that high-engagement trainings are more effective than low-engagement ones (Robson *et al.*, 2010).

Sacks *et al.* (2013) focused on improving workers' hazard identification skills, rather than on prescriptive instruction of classic methods to perform specific tasks by conducting a 'between-participant' experiment with a comparison of conventional safety training methods and immersive virtual reality safety training methods. They found that virtual environment provides its users with a better understanding on safety issues (Sacks *et al.*, 2013).

2.3. New Safety Training Methods in the Construction Industry

It is crucial to find out different ways of learning for workers' training especially active means of learning with developing information technology (Bhoir and Esmaeili, 2015). In terms of workers' learning, the least-educated group of people get less access to learn new things at work. The proportion of workers who report developing their skills at

work varies between 72% and 84% in the United States and the EU however, the proportions declines to 57% in Turkey (Eurofound and ILO, 2019). In this regard, researchers have questioned the effectiveness of traditional safety training programs and focused on how to develop the learning quality of safety training. Guo *et al.* (2012) have summarized the main reasons for the high accident rate in construction industry as the complex nature of construction which prevents people from predicting potential safety problems and the safety training methods where trainees can only listen and watch without a participation. Training methods are evolving by time and safety trainings has begun to be done by computers (Van Buren and Erskine, 2002) and numerous approaches integrated with new technology has been introduced by researchers.

2.3.1. Reality Technologies

Safety training tools have begun to be integrated with reality technologies to enhance the current educational strategy since virtual learning environment, provides its users with interactive teaching patterns and improve its users' ability of exploring new concepts (Pan *et al.*, 2006). Reality technologies can be summarized in the continuum developed by Milgram and Kishino in 1994 as shown in Figure 2.2. Milgram and Kishino (1994), proposed a Reality-Virtuality Continuum ranging from a real environment to virtual environments in which the degree of computer-generated stimuli increases going to the right end of the continuum. Real Environments (RE) which is stated in the left end of the continuum cover the reality by providing views of real scenes directly or indirectly, i.e. through video display. On the other hand, in the right end of the continuum, Virtual Environments (VE) are completely computer-generated environments in which physically non-existing objects are "displayed" on a device and its users can explore and interact in real-time through a technological interface (Bhoir and Esmaeili, 2015). Virtual reality (VR) is a technology that involves in computers, software and hardware to simulate a real or imaginary environment for its users (Sacks *et al.*, 2013). On the other hand, an immersive virtual environment provides its users with a sense of being within this computer-generated environment by decreasing the users' perception of the real environment (Sacks *et al.*, 2013).

Within this continuum, the remaining realities placed between the real environment and virtual environments are termed as Mixed Reality (MR) environments in which real and

virtual objects were merged. (Milgram and Kishino, 1994). MR consists of augmented reality and augmented virtuality (Milgram and Kishino, 1994).

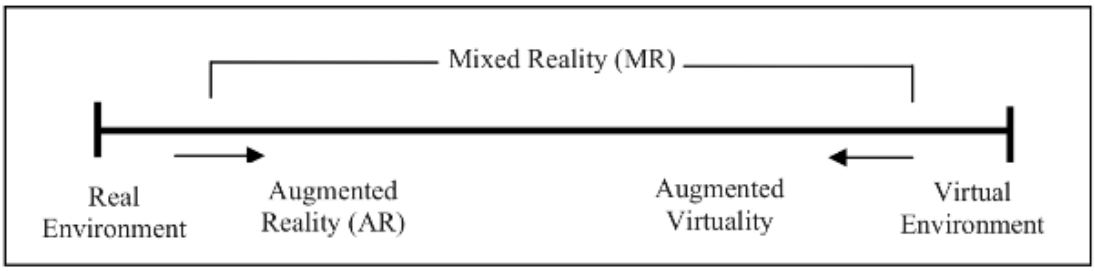

Figure 2.2. Milgram Reality-Virtuality Continuum (Milgram and Kishino, 1994)

An immersive virtual environment (IVE) is a computer-generated environment that provides its users with a feeling of being within its environment by engaging with users' senses and reducing the perception of the real environment (Sacks *et al.*, 2013). Various game engines are now available with different programming levels to adapt for virtual reality environments such as 3DVIA Virtools and Unity which are user friendly (Li *et al.*, 2012). Guo *et al.* (2012) have presented a game technology-based safety training approach as a potential in solving safety problems of construction plant operations. Mohd *et al.* (2019) have proposed a serious game approach to be applied in safety training tools which includes a visual training with a real case scenario that helps workers to achieve specific subjects in a safe and entertaining way. Zhao and Ye (2012) also used virtual environments in the development of 3D online computer games to increase electrical hazards awareness in the U.S. construction industry.

2.3.2. Virtual Reality Safety Training Tools in Construction Industry

Virtual reality technologies have been adopted in many sectors for the training purposes (Irizarry and Abraham, 2005). Training tools in virtual reality have been more feasible in terms of workers' development in learning skills and concentration comparing to the conventional training methods (Sacks *et al.*, 2013). Among the new technologies, virtual reality environments are becoming more widely used for educational purposes (Bhoir and Esmaeili, 2015). Safety trainings developed with virtual reality technology provides its users with better understanding of safety rules, standards and regulations by creating 3D virtual

images and experiences for the users in a safe and controlled environment (Xie *et al.*, 2006; Bhoir and Esmaeili, 2015). Virtual Environments (VE) have been explored by many researchers as an effective way of safety learning for workers (Kassem *et al.*, 2017). Many studies have been done by using virtual environments in the development of safety tools such as the 3D interactive safety assessment (Li *et al.*, 2012), a virtual construction environment for steel erection (Irizarry and Abraham, 2005), a multi-user virtual reality system with motion tracking function (Shi *et al.*, 2018) and a VR-based safety training system for construction workers' electrical hazards awareness enhancement (Zhao and Ye, 2012).

Training with traditional methods (by lectures or demonstrations) which is categorized as low engagement has been found to have less effective comparing with more engaging training methods (Burke *et al.*, 2006). Zhao and Ye (2012) have pointed out various features of the trainings developed in virtual environments by comparing the traditional classroom-based trainings. First, the users have a chance to experience the hazards in the risk-free environment and find the correct way to accomplish the task which provides interactivity to improve their safety skills (Zhao and Ye, 2012). The trainees can "physically interact" with the virtual environment and experience different working conditions with minimum risk to the trainee (Irizarry and Abraham, 2005). Secondly, the users' performance can be tracked and evaluated through the application which lowers the need for direct trainee observation (Zhao and Ye, 2012). Lastly, the trainings in virtual environments provides its users with convenience of location and time unlikely the traditional classroom-based trainings where the coordination and availability of the trainees and the training providers needs to be scheduled beforehand (Zhao and Ye, 2012).

Virtual reality also eliminates the restrictions due to weather conditions and other external obstacles that may make trainings difficult to conduct in real environment (Bhoir and Esmaeili, 2015). Shi *et al.* (2018) used virtual environment in order to simulate hazardous situations in their experiment because hazardous cases cannot simulate in real-world scenarios easily. Another feature that VR offers is that multiple users can interact in the same simulated environment (Bhoir and Esmaeili, 2015). This feature enables users to evaluate how the others’ decision on health and safety can affect the overall result and how to work with others collaboratively. In this regard, Guo *et al.* (2012) developed a serious

game platform which enables trainees to operate a construction plant and to collaborate in virtual multiuser environment. Li *et al.* (2012) also used the multi-user feature of VR in the safety training game engine system for tower crane dismantlement. This system provides the users with a risk-free environment to follow the dismantling procedure and to collaborate with other users.

Traditional classroom trainings are lack of drawing attention of the participants. Sacks *et al.* (2013) has drawn attention to three important conclusion in their study: (i) virtual reality has been found to be suitable for presenting hazards to trainees realistically without posing any risk to their safety, (ii) virtual reality safety training has captured trainees' attention better compared with conventional classroom trainings, (iii) virtual reality safety training tool has provided trainees with a control over the environment therefore it has also reinforced their learning.

Messner and Horman (2003) observed the potential benefits of using virtual reality technology in educating construction engineering by conducting an experiment with 4D CAD tools for students. Their study revealed that students can perform better results in construction planning by using advanced visualization tools. Trainings developed in virtual environments can also capture the participants' attention and concentration (Shi *et al.*, 2018).

Students have a chance to observe the construction activities only in site visits and it generally ends up with a passively observing the process done by others. However, one of the most effective way of learning is “learn by doing” (Xie *et al.*, 2006). Therefore, students need to participate in the activities and take an active role for better understanding of construction. In this regard, virtual reality tools help students take place in construction projects in virtual environments and experience how their decisions will have an impact on the virtual project (Messner and Horman, 2003).

During the development stage of a new safety training tool, typically there is back and forth process since it requires many adjustments and revisions. However, scheduling the experiments with the construction workers requires time and coordination which in turn leads a considerable time. Therefore, before scheduling the experiments, the developed safety training tool should be end-product and tested with an experiment group beforehand.

In this regard, students might be a suitable and efficient way of target group in developing of the safety training tool. There is a gap in the literature on this issue that students can be used as surrogates in developing the safety training tools in construction industry. Considering this, the objective of this thesis is to investigate whether students can be used as surrogates in evaluating the safety training tools.

2.4. Appropriateness of the Use of Students as Surrogates

Research areas in many sectors use students as surrogates in the subject selection of their research. However, the appropriateness of this practice has been questioning and the results are inconclusive (Trottier and Gordon, 2018). In this regard, more research on this issue should be done in order to find a comprehensive result at some degree (Trottier and Gordon, 2018). This practice has been commonly adopted in managerial studies, specifically. Most studies in behavioral accounting and management information systems employ students as surrogates for the real-world people (Ashton and Kramer, 1980; Hughes and Gibson, 1991).

The use of students as surrogates is a quick and convenient source of information in terms of time, distance, and money considerations (Shuptrine, 1975). Abdel-Khalik (1974) demonstrated two reasons why researchers tend to use students as surrogates for industry managers: the lack of cooperation by managers in experimental studies, and the high cost of scheduling field experiments. Additionally, it is quite time-consuming to conduct the experiment with the target group because adequate number of people and convenient conditions for the experiment cannot be found easily. Especially, if the target group is professionals or workers in the specific sector, the coordination might be challenging for the experiments due to their tight schedules. Even if the researchers schedule an experiment with the target group, the experiment might need to be repeated in case of the adjustment in the structure of the experiment. There is a long back and forth process in the development of the tool that the researchers are working on. Considering all these reasons, students can be deemed as convenient resources to be used as surrogates for the target group in many research areas.

Despite the reasons for the common use of students as surrogates, researchers are divided as proponents and opponents in the sense of the appropriateness. Are students good surrogates for the real-world professionals or any kind of subject group? The answer of this question depends on various variables and conditions. Apart from the results, the robust conclusion is hard to be made that students are always or never appropriate surrogates for managers (Trottier and Gordon, 2018).

2.4.1. Use of Students as Surrogates in Managerial Studies

Use of students as surrogates is a common phenomenon in managerial studies as well. They also questioned the appropriateness of this phenomenon (e.g. Abdel-Khalik, 1974; Ashton and Kramer, 1980; Elliott *et al.*, 2007; Mortensen *et al.*, 2012; Trottier and Gordon, 2018). Abdel-Khalik (1974) conducted an experiment on 40 lending cases with MBA students compared to managers and found that students are not suitable surrogates for managers. On the other hand, Ashton and Kramer (1980) compared the performances of students and auditors on internal control judgment cases and found that students are suitable surrogates for auditors. Elliott *et al.* (2007) also conducted a study with MBA students and non-professional investors and found that MBA students who took financial analysis course are good surrogates for non-professional investors. Mortensen *et al.* (2012) conducted similar experiment with auditing students and auditors based on ten accounting judgements and found that students might be appropriate surrogates for accounting professionals in decision-making practices.

Trottier and Gordon (2018) have questioned when students can be convenient subject group for managers and made an extensive literature review on this issue. Trottier and Gordon (2018) demonstrated the proponents and opponents in a brief summary table derived from their study as shown in Table 2.1. They made an extensive literature review based on the context and the surrogacy conclusion. This table interprets that the appropriateness of the use of students as surrogates is depending on the context. Students are not good surrogates in behavioral studies however, they are good surrogates when they approximate the target group (Trottier and Gordon, 2018).

Table 2.1. Summary of Literature Review done by Trottier and Gordon, 2018.

Article authors, date	**Dimension/Mechanism**	**Surrogacy conclusion**
Alpert, 1967	Knowledge; Attitudes	Poor surrogates
Enis *et al.*, 1972	Perspective; Attitudes	Inconclusive
Copeland *et al.*, 1973	Attitudes	Poor surrogates
Shuptrine, 1975	Consumer attitudes	Inconclusive
Abdel-Khalik, 1974	Decision making	Poor surrogates
Ashton and Kramer, 1980	Decision making	Adequate surrogates
Weber, 1980	Decision making; recall	Inadequate
Remus, 1986	Decision making	Good surrogates
Libby and Frederick, 1990	Decision making	Inadequate
Hughes and Gibson, 1991	Decision making	Inadequate
Frederick, 1991	Recall	Inadequate
Tubbs, 1992	Knowledge	Inadequate
Houghton and Hronsky, 1993	Recognition; Measuring	Good surrogates in recognition; not meaning
Nelson, 1993	Understanding and knowledge	Inadequate
James and Sonner, 2001	Reactions, perceptions	Good surrogates when similar age
Bean and D'Aquila, 2003	Decision making	Unsuitable
Hoffman *et al.*, 2003	Decision making	Good surrogates
Chang and Ho, 2004	Decision making	Unsuitable
Fehr and List, 2004	Decision making	Unsuitable
Liyanarachchi and Milne, 2005	Decision making	Good surrogates
Elliott *et al.*, 2007	Decision making	Good surrogates
Mortensen *et al.*, 2012	Decision making	Good surrogates
Trottier and Gordon, 2016	Decision making	Good surrogates

Abdel-Khalik (1974) conducted an experiment on 40 lending cases with MBA students compared to managers and found that students are not suitable surrogates for managers. On the other hand, Ashton and Kramer (1980) compared the performances of students and auditors on internal control judgment cases and found that students are suitable surrogates for auditors. Elliott *et al.* (2007) also conducted a study with MBA students and non-professional investors and found that MBA students who took financial analysis course are good surrogates for non-professional investors. Mortensen *et al.* (2012) conducted

similar experiment with auditing students and auditors based on ten accounting judgements and found that students might be appropriate surrogates for accounting professionals in decision-making practices.

Since the results on this issue produced mixed conclusions, there is no clear conclusion that students are always or never convenient surrogates for managers (Ashton and Kramer, 1980). Trottier and Gordon (2018) conducted an experiment with students and senior managers in financial reporting and found that although student and manager responses are significantly different statistically, both groups end up with the same conclusion in the scope of the experiment. Trottier and Gordon (2018) did not focus on the statistical results of the responses in the experiment because it would lead to incorrect conclusions without considering how the differences affect the conclusion. Therefore, the differences between students and the target group does not indicate that students are not suitable surrogates (Trottier and Gordon, 2018).

2.4.2. Use of Students as Surrogates in the Virtual Safety Training Tools

Using students as surrogates is a common phenomenon among academic studies. Many research areas used student group as surrogates for workers in the health and safety subjects because of the above-mentioned reasons. They mainly do not focus on whether students are good surrogates for workers. For example, Becker and Morawetz (2004) carried out a survey with students to evaluate the worker activities before and after training regarding chemical spills and other chemical exposures. Irizarry and Abraham (2005) conducted two experiments for the VR safety application in the steel erection process. They used students as a subject group for the first experiment and ironworkers as a subject group for the second experiment (Irizarry and Abraham, 2005). Shi *et al.* (2018) developed a multi-user virtual reality system to observe the effect of social influence on workers' safety behaviors, however, they conducted the experiments by using students as surrogates for the construction workers. Xie *et al.* (2006) also used a virtual reality safety training system to evaluate the behavior impact of the tool on trainees. Perlman *et al.* (2014) structured the experiments with both students and professionals in the construction industry as surrogates for the construction workers. They compared the responses of each group and drew remarkable

conclusions however they also noted that the direct evaluation of workers' perception would be important for the future (Perlman *et al.*, 2014).

Although there are many research areas on safety training tools used students as subject group in the experiments, there is a gap in questioning whether students are good surrogates for construction workers in the development of virtual safety training tools. In this regard, the objective of this thesis is to fill this gap in the literature by investigating the appropriateness of students as surrogates for the construction workers and for engineers in the development of virtual safety training tool applications.

3. METHODOLOGY

Methodology of this thesis is explained under four sections of this chapter. In the first section, hypotheses are presented by providing similar studies in the literature. After defining the hypotheses, virtual safety training tool named V-SAFE v.2 is introduced in the second section by identifying the risk sources in the tool and by explaining its training and testing modules. Experimental settings are explained by covering the selection of subject groups, experiment flow and experiment materials in third section. Lastly, appropriate analysis methods to test the hypotheses are presented based on the similar studies in the literature.

3.1. Hypothesis Development

The objective of this part is to develop a set of hypotheses based on the aim and objectives of this thesis. In this regard, the literature review and related methodologies in previous studies have been presented and a set of hypotheses have been developed accordingly.

Five hypotheses are stated in which cover the aim and objectives of this thesis. First three hypotheses are determined to test if there is a significant difference in the participants' knowledge regarding basic scaffolding tasks and related health and safety implementations after receiving virtual safety training tool for students, construction workers and engineers, separately. Similar study takes place in the experiment conducted by Sacks *et al.* (2013). They prepared a set of experiments to compare the subject group's safety test scores collected just prior to the training with the scores collected after the training (Sacks *et al.*, 2013). Moreover, Pedro *et al.* (2016) assessed the participants' performances through a paper-based test in order to observe the difference between the education scores of the participants before and after taking a developed Virtual Construction Safety Education System framework. Reid and Dawson (2001) conducted a study to compare the performance difference between the older and younger subject groups on 12-hour shift rotation. The aim of their study is to evaluate if there was a meaningful difference in performance between the shifts for both the older and younger subject groups (Reid and Dawson, 2001).

Considering all these relevant studies, first three hypotheses are established in order to investigate the effectiveness of virtual safety training tools in terms of enhancing the safety knowledge and skills. Hypotheses 1,2 and 3 are developed based on this aim, however, the subject group is different for each hypothesis. The null and alternative hypotheses for hypothesis 1, 2 and 3 are listed as the following:

Hypothesis 1:

H_0: There is not a significant difference between the paper-based test scores of students before and after receiving virtual safety training.

H_1: There is a significant difference between the paper-based test scores of students before and after receiving virtual safety training.

Hypothesis 2:

H_0: There is not a significant difference between the paper-based test scores of construction workers before and after receiving virtual safety training.

H_1: There is a significant difference between the paper-based test scores of construction workers before and after receiving virtual safety training.

Hypothesis 3:

H_0: There is not a significant difference between the paper-based test scores of engineers before and after receiving virtual safety training.

H_1: There is a significant difference between the paper-based test scores of engineers before and after receiving virtual safety training.

The use of students as surrogates have been investigated in many research areas. Hughes and Gibson (1991) conducted an experiment in their study to compare the results of students and managers on the use of a decision support system generator by using a training

program. Their aim was to evaluate if students and managers made similar responses before and after the training in the use of the generator (Hughes and Gibson, 1991). Perlman *et al.* (2014) adopted a research method to make a comparison of the hazard identification and assessment skills of three subject groups: construction superintendents, civil engineering students and safety directors. As a result of related literature review, hypotheses are stated as the following:

Hypothesis 4:

> *H_0: There is not a significant difference between the paper-based test results of the construction workers and students after receiving virtual safety training.*
>
> *H_1: There is a significant difference between the paper-based test results of the construction workers and students after receiving virtual safety training.*

Hypothesis 5:

> *H_0: There is not a significant difference between the paper-based test results of the engineers and students after receiving virtual safety training.*
>
> *H_1: There is a significant difference between the paper-based test results of the engineers and students after receiving virtual safety training.*

Hypothesis 4 and 5 are determined to seek if the safety performances of students on the paper-based test results are similar with construction workers and engineers respectively.

3.2. Introduction to V-SAFE v.2 Training Tool

Newly developed virtual environment-based safety training tool named V-SAFE v.2 is adopted in the experiments to test the hypotheses. V-SAFE v.2 training tool has been previously introduced in TUBITAK project within Grant No. 315M186 "*Development of Worker Health and Safety Simulation Tool in the Virtual Environment*" (Çomu, 2018). One of the main objectives of the project was to minimize occupational accidents occurring

during the installation of H-type scaffolding through the virtual environment-based safety training tool. Besides introducing the safe working procedures, the tool also supports the improvement of hazard identification and collaboration skills, which are the crucial health and safety concepts in the construction sites. V-SAFE v.2 has been adopted in this thesis to evaluate the appropriateness of students using as surrogates for the construction workers and engineers in the development of virtual environment-based safety training tools. While doing this, the effectiveness of the tool on these three subject groups' learning is also evaluated. The newly developed virtual safety training tool includes the appropriate visuals of the equipment used in the installation of H type scaffolding and the representative players that users can have a control in their movements to complete the tasks in the virtual environment. There are three roles in the tool that maximum three users can involve in. This enables users with the single and multi-user environments so that the users can use the tool as a team or on their own depending on their choice. If they choose to use the tool on their own, the autonomous agents take place for the remaining two roles, therefore they can still use the tool with a team. Risk sources and the visuals in the tool and the scenario development are explained in the following subtitles respectively.

3.2.1. Identification of Risk Sources and Development of the Visuals

Identification of Risk Factors: Risk factors and classification of occupational accidents in construction sites are important in terms of decreasing death and injury rate. One of the most common methods used for risk determination is to identify risks by considering past accidents. Additionally, local factors play an important role in determining risk factors. Therefore, accidents occurred in Turkey were considered for this project in order to determine the risk factors using a more realistic approach.

Considering the high accident rate resulting in death and injury in the construction sector, it is vital to determine the risk factors in trainings and occupational accidents through virtual reality. Thus, risk factors can be identified accurately. Risk identification contributes to the creation of more reliable business environments. The types of accidents in the construction site were generally determined by considering the data of the accidents that occurred in the past (Gürcanlı and Müngen, 2013). Additionally, the accidents taking place in the virtual environment-based safety training tool developed in the previous BAP Project

under Grant No. 7902 has also considered. Gürcanlı and Müngen (2013) examined the accident data that occurred during scaffold formwork installation and found that: loss of balance on the scaffold, breaking of planks, collapse of scaffold, scaffold mantling and dismantling, breaking of horse scaffold equipment are associated with the falls from the scaffold. This data is reviewed to determine the risk factors to be integrated into V-SAFE v.2.

The visuals in 3D Modelling: V-SAFE v.2 presents the visuals of materials and equipment used in H type scaffolding process; therefore, their 3D models should be developed accordingly. At this point, it is crucial that 3D models of all materials and equipment are designed as real as possible for enhancing the effectiveness of the safety training simulation tool. The sense of reality in virtual environments increases the participants' attention and concentration (Shi *et al.*, 2018). Before visualizing unsafe site situations, an appropriate site environment should be determined. Therefore, relevant materials and equipment for the installation of H-type scaffolding (e.g. scaffold's crossing profiles, adjustment pins, wooden railings, etc.) have been determined accordingly. The final list of material drawings has been prepared, and the agreed products have been supplied to the construction company. The 3D models of the scaffolding and concrete formwork materials have been developed in the most realistic way. Thus, it is aimed to provide a more comprehensible scaffolding and concrete formwork installation training in order to ensure worker safety. Thus, the materials to be used at each stage, the task sequence and how the workers would move together have been approved by the project team. These approved materials and task sequence were integrated in the virtual environment appropriately. In this regard, a virtual demonstration of the training including the agreed materials and task sequence have been requested from the simulation developer.

3.2.2. Scenario Development

Once the risk factors and necessary materials have been identified in the virtual-based environment, the scenario has been developed in a way that it should cover the basic steps of the installation of H-type scaffolding in line with the health and safety requirements. In each step, the tasks have been structured so that there is a meaningful flow between the tasks. This flow between the steps have been developed and integrated in the tool. The

scenario has been developed to highlight the tasks in which the occupational accidents could occur in case of unsafe acts and the instructions of H-type scaffolding and concrete formwork installation. Therefore, in the scenario, the sections carrying risks for the workers during the H-type scaffolding and concrete formwork installations have been highlighted. In other words, the purpose of the developed tool is to successfully carry out the tasks in line with the safety regulations by the participants.

V-SAFE v.2 consists of two modules: (i) the training module which is designed to cover the basic task sequence of H type scaffolding and related health and safety implementations and; (ii) the testing module in which is designed to evaluate the performances of the participants which is their knowledge that they gained from the training module of V-SAFE v.2. During the training module, participants are provided with information boxes covering safety instructions that they can follow throughout the training without any accident.

Training and Testing Modules in V-SAFE v.2: The scenario of scaffolding installation tasks in line with health and safety implementations has been determined as two modules: (i) training module and (ii) testing module. In both modules, the scenario flow remains same, however, the participants are directed by the information boxes and flashing objects only in the training module. On the other hand, in the testing module, the participants are expected to remember the sequence of the works and the basic health and safety requirements. Therefore, they should perform the tasks based on what they have learnt in the training module. In other words, if the participants make a mistake in the testing module, they can experience the potential results of the failure. For example, if a representative player forgets to lock the pins of the crossings, the H scaffold would collapse before moving to the next step.

The training module of the tool is designed to provide the basic task sequence of H type scaffolding and related health and safety implementations. In other words, the participants can follow the information boxes which guides participants what to do in every step. Therefore, it is not possible for the participants to make mistakes during the training module. In V-SAFE v.2, the importance of using personal protection equipment (PPE) are emphasized as well. Every step is described and instructed with an information box

throughout the training module. In addition to information boxes, participants are directed to the flashing objects which means that they should go to the directed place. Hence, participants can follow these information boxes and direct to the flashing objects with point and click. As the participants move to the next step with point and click mechanism, they can direct their representative players with a single command. In this way, the loss of time is prevented, and the training flow progresses more rapidly. The task flow in the training module is summarized step by step below.

1. Firstly, the participants click on the flashing equipment storage to select personal protective equipment that should be used at the construction site (Figure 3.1).

Figure 3.1. Beginning of V-SAFE v.2 Training Module.

2. After clicking the flashing equipment storage, the representative players are automatically directed to that area. The equipment storage includes the personal protection equipment (hard hat, goggles, belt and glove) that the participants must take before working at the construction site. The participants choose the required personal protective equipment by clicking on the items (Figure 3.2).

Figure 3.2. Personal Protection Equipment List in the Equipment Storage.

3. As the participants select each personal protective equipment, the selected equipment is shown on the representative players (Figure 3.3).

Figure 3.3. Representative Player with the Selected Personal Protection Equipment.

4. After the selection of personal protection equipment, it is necessary to go out from the material storage to carry the materials for the formwork as it is also written in the information box (Figure 3.4).

Figure 3.4. Flashing Object Pointing the Material Storage Area.

5. Since one worker cannot carry a load of more than 25 kg according to the occupational safety rules, another worker should be called for help (Figure 3.5). To do this, the participants are directed to click on the radio icon on the left tab of the simulation interface. Thus, by selecting the message to be sent from the menu, communication with other workers can be made properly (Figure 3.6).

Figure 3.5. Warning Box to Call Workers for Help.

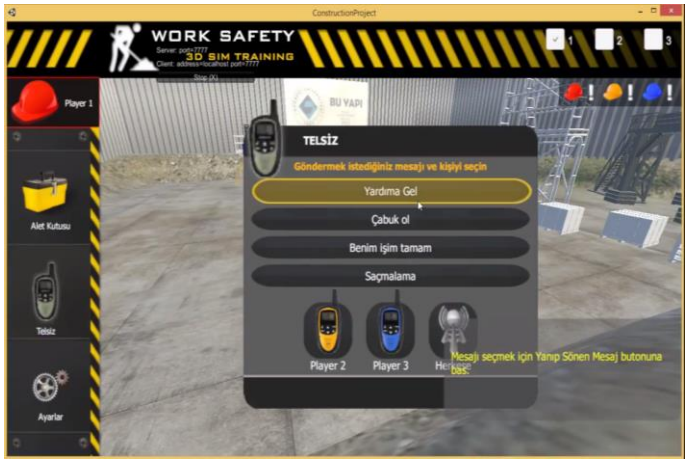

Figure 3.6. Radio Icon to Call Help.

6. After the participants take the materials for column formwork and carry them to the specified area, the steps related to the column formwork system are shown in a summarized manner automatically. After the column formwork installations are completed, the H type scaffolding installation process is started.

7. The materials required for the H type scaffold to support the underfloor formwork system are selected by clicking the toolbox on the left tab of the simulation interface (Figure 3.7). Not all materials can be selected at once. Participants are asked to select the appropriate material in line with the sequence of the tasks. So, participants should click on the toolbox to select the required material in each step.

Figure 3.7. Material List in the Toolbox.

8. Before the installation of H type scaffolding, the base plates must be placed on the ground to transfer the load to be come from the adjustable jacks (Figure 3.8). Therefore, the participants select the base plates by clicking the visual on the toolbox.

Figure 3.8. Base Plates on the Ground.

9. Then, the information box directs the participants to select the adjustable jacks to be placed on the base plates. The adjustable jacks are placed by clicking and pointing the specified place (Figure 3.9).

Figure 3.9. Adjustable Jacks on the Base Plates.

10. H type scaffoldings are placed on the adjustable jacks in order to install the H type scaffolding system (Figure 3.10).

Figure 3.10. H Type Scaffoldings on the Adjustable Jacks.

11. To secure the H scaffold, the diagonals must be fitted, and the pins locked as it is instructed in the warning box (Figure 3.11). Similar to the previous steps, participants click on the pin to lock and stabilize the system. The warning box shown in Figure 3.11 include an information that they should lock the pins before moving to the next step.

Figure 3.11. Warning Box to Lock the Pin.

12. In the next step, the planks of appropriate size should be placed on the scaffold and railings should be installed for the workers to work safely on the cat path (Figure 3.12).

Figure 3.12. Representative Player Standing on the Planks.

13. Once the H type scaffolding system has been installed, workers must attach a parachute seat belt to the lifeline to prevent falling from a height in order to work safely on the cat path and complete the training module (Figure 3.13).

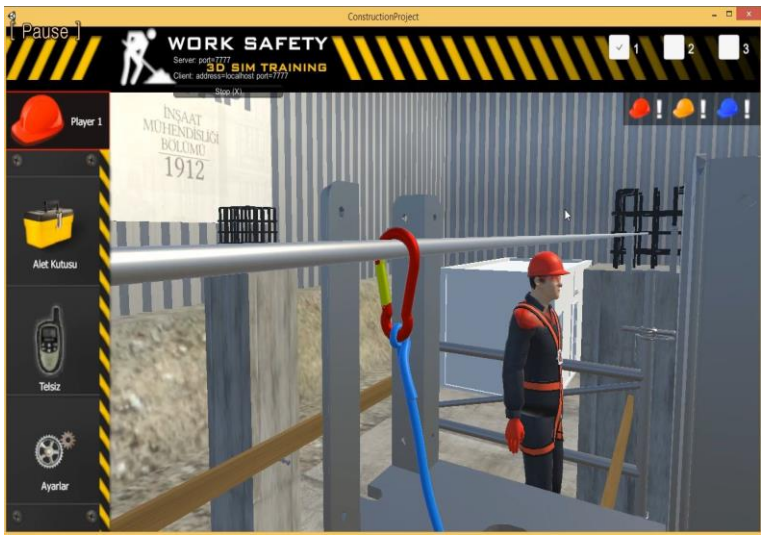

Figure 3.13. Parachute Seat Belt Attached to the Lifeline.

The testing module is designed to evaluate the performances of the participants which depend on the knowledge that they gained in the training module. The purpose of the testing module is to measure the effectiveness of the simulation training on the performance of the worker. Unlikely in the training module, there is not any information boxes to guide the participants in this module. Therefore, the testing module in V-SAFE v.2 training tool allows the participants to make their own decisions regarding the safety applications for H type scaffolding based on what they have learnt in the training module. So, if the participants make a mistake in this module, they will experience the accidents resulting from their mistake. For example, if they forget to take a helmet, an object will collapse on their representative player's head, the player gets injured and the game is over. Moreover, if they choose to use the tool in multiuser platform and one player makes a mistake, all players experience the consequences of that mistake. For example, if one participant forgets to lock the pin in the crossing of the H type scaffold, the scaffold will collapse and therefore all players on the scaffold will fall. In this way, the simulation of the scaffolding and formwork installation will be designed as close to reality as possible and the participants will be able to experience real accident scenarios. Considering that one of the most important problems in the construction site is that workers do not have sufficient awareness and knowledge about accidents, it is aimed to contribute to the creation of safer workplaces by increasing the safety awareness of the workers at the construction site.

In the testing module, each participant starts the simulation with a total of 100 points. Score is reduced in accordance with the mistake or missing activity that the participants might make. In other words, a high score indicates that the participant completed the activities in line with the safety rules. In addition, another important feature of the testing module is that the participants will experience results of their unsafe acts.

Autonomous Agents: Autonomous agents in modern computer games are called NPC (Non-Playable Character). Autonomous agents have been used in various training simulation trainings (e.g. Dickinson *et al.*, 2011; Edward *et al.*, 2008). Dickinson *et al.* (2011) used autonomous agents in the pit digging simulation to increase realism. The autonomous agents do not have the ability to decide in the background and do the same job continuously (Dickinson *et al.*, 2011). Edward *et al.* (2008) modelled autonomous agents as factors which have a certain artificial intelligence in the factory showing the behaviors that may cause

accidents. There are three roles in both training and testing modules of V-SAFE v.2. The virtual environment-based training simulation tool is designed to train participants about basic tasks in scaffolding installation and relative health and safety issues.

The tool provides its users with two different options; (i) multi-user environment; (ii) single-user environment. The participant can select the environment (either multi-user of single user) before starting the simulation training (Figure 3.14 and Figure 3.15). In multi-user environment, all roles are managed by different users and therefore there is no need for autonomous agents. In multi-user environment, three users can follow the steps and perform tasks collaboratively. If one of them makes a mistake (i.e. not locking the pins in the crossing of H type scaffolding), all users are exposed to the consequence of the mistake.

In single user environment, the user performs the tasks and other roles are managed by autonomous agents. In the single-user environment, a user can perform tasks and the other two roles are governed by the autonomous agents. Therefore, the user can experience performing tasks collaboratively with autonomous agents. Meaning that, both environments provide its users with the experience of performing tasks collaboratively. In this study, all participants took V-SAFE v.2 in a single user environment. Since the scenario is composed of three different roles, the participants can distinguish their representative players according to the helmet color they choose (red, yellow, blue).

In single user environment, the tool provides its users with the autonomous agents in two options: (i) autonomous agents performing tasks without a mistake; (ii) autonomous agents performing tasks with unsafe acts. In this study, however, the autonomous agents performing tasks without a mistake were selected. Since the aim is to assess the participants' performance only.

Figure 3.14. Selecting V-SAFE v.2 Training Module with Autonomous Agents.

Figure 3.15. Selecting V-SAFE v.2 Testing Module with Autonomous Agents.

3.3. Experimental Settings

The experimental setting covers the materials and subject groups to perform the experiments in an effective manner. The primary aim of the set of experiments is to observe if student subjects can be used as surrogates for the construction workers and engineers in the development of the virtual environment-based safety learning tools. By doing this, the effectiveness of V-SAFE v.2 on learning of the three subject groups can be also evaluated accordingly. A set of experiments are designed to measure first the effectiveness of V-SAFE

v.2 on the subject groups and then to evaluate the appropriateness of using students as surrogates for the construction workers and engineers in this type of safety training related studies.

3.3.1. Selection of Subject Groups

The subject groups were categorized into three groups: (i) students; (ii) construction workers; (iii) engineers. The experiments were performed with 20 construction workers, 20 undergraduate civil engineering students and 20 engineers. These three groups were compared between groups and within groups in terms of their performances on V-SAFE v.2 training tool by considering their backgrounds.

Student Group: The student group consists of 20 senior civil engineering students in Boğaziçi University. Experiments with student subject group were conducted in Boğaziçi University Civil Engineering Department.

Construction Worker Group: The appropriate construction site environment was found to conduct the experiments with construction workers. To conduct the experiments with the construction workers at the construction site is a realistic way to test V-SAFE v.2 training tool. By doing this, the tool was tested directly by the target group which were construction workers. Therefore, the effectiveness of the tool can be observed more accurately by evaluating their safety performances in the tool. Arrangements were made regarding the date and location of the experiments. Since there was a time limitation at the construction site, the experiments were conducted with 20 construction workers in two technical offices at the construction sites.

Engineer Group: The engineer group consists of 20 engineers from mechanical, civil and environmental engineering background in the construction companies. The half of the experiments were conducted in the construction sites where the worker data was collected. Rest of the experiments were administered in a construction consultancy company.

3.3.2. Experiment Flow

Experimental setting is designed so as conduct the experiments in an effective manner. Experimental setting consists of three main stages as summarized in Figure 3.16.

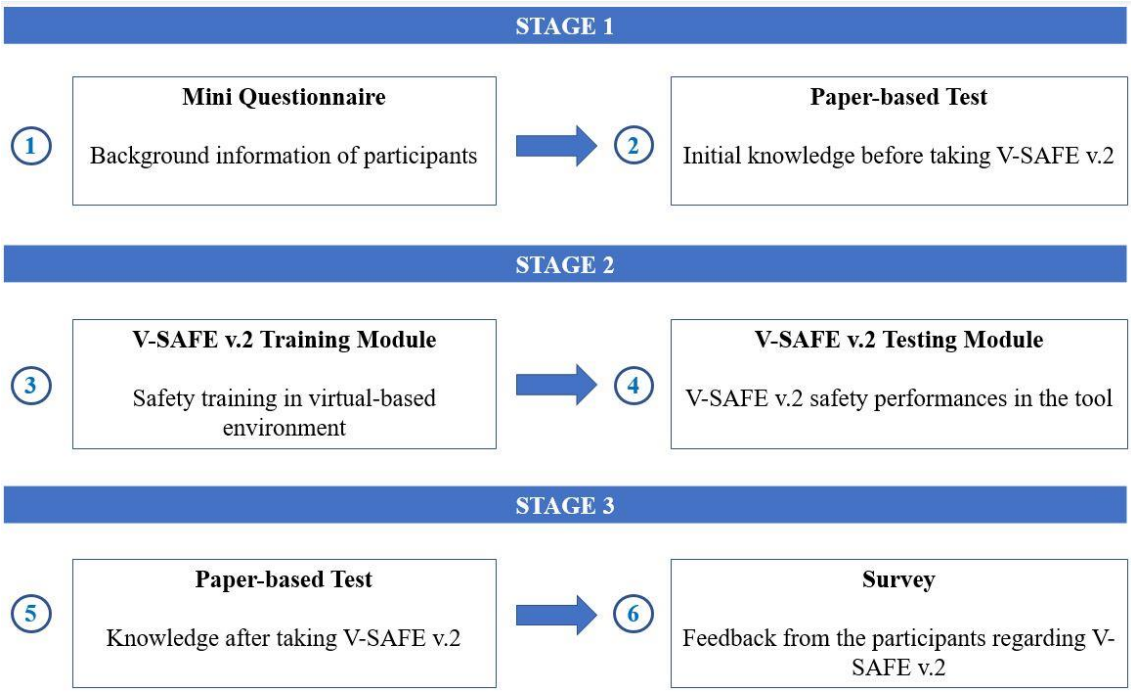

Figure 3.16. Flow Chart of Experimental Procedure.

V-SAFE v.2 has been verbally introduced to participants before starting the experiments. In Figure 3.16, the first stage of the experiment involves a brief written information about V-SAFE v.2. Mini questionnaire was administered to gather the demographic information of participants (e.g. age, work experience, their role in the construction industry etc.). After filling mini questionnaire, participants were asked to take the paper-based test. The paper-based test consists of 10 questions covering basic tasks in scaffolding installation and related health and safety implementations covered in V-SAFE v.2. This paper-based test was administered before receiving V-SAFE v.2 training module in order to see the participants' initial knowledge on these subjects. In the second stage, participants took V-SAFE v.2. training module. After completing the training module, they were asked to move on to the testing module of V-SAFE v.2. At the end of the testing module, summary table showing their safety performances appears on the computer screen.

In this way, safety performance on V-SAFE v.2 could be recorded for each participant. After receiving V-SAFE v.2, participants were asked to retake the same paper-based test as they have completed at the first stage. In this way, they could reflect what they have learnt in the training tool by answering the same questions. In the third stage, the participants were asked to retake the same paper-based test at the second stage. Lastly, participants were asked to complete the 13-question survey based on the Kirkpatrick model. Therefore, feedbacks from participants regarding V-SAFE v.2 could be collected through survey. The details of the experiment materials and the analysis methods are explained in the following subtitles.

3.3.3. Experiment Materials

The experiment materials include a mini questionnaire, paper-based test, V-SAFE v.2 training and testing modules and survey respectively. The experiment materials are explained as the following: mini questionnaire, paper-based test and survey form.

In the beginning of mini questionnaire, a consent form was attached which was approved by the Ethics Committee of Boğaziçi University. The form includes a signature part indicating that the participants voluntarily participated in the experiment and the personal information stated in this form for each participant can be used in this project with the consent of a participant. Mini questionnaire form is administered to collect general information about the participants' backgrounds. Killingsworth *et al.* (2014) adopted similar type of survey to gather information about the participants' backgrounds in their research protocol. There were six questions including what role they were working at the site, how many years they had been working at the site and what kind of computer games they had played before. Since there are three subject groups in the experiments, some of the questions are adjusted for each group while the consent form remained same. For the student subject group, since they did not have work experience, the work-related questions were adjusted as whether they had an internship experience at the construction site before and what grade they were in. For the engineers, they were additionally asked if they had any health and safety implementation experience. The rest of the questions remained same. In addition to these questions, whether the participants had played computer games in virtual environment before were also asked. A table was placed in this question and the names of the generally known games were written and asked how often they played. They were also asked if they

had ever played construction simulation games (Construction Simulation, Truck Simulator) before. Thus, an evaluation can be made according to V-SAFE v.2 test results among the participants who have played construction simulation games before and have not played. The Consent form given to each participant was completed and general information of the participants was obtained.

Paper-based tests are administered to measure the participants' initial knowledge of occupational health and safety applications and basic scaffolding installation steps before receiving V-SAFE v.2 training module. The test involved ten questions covering the basic applications in H type scaffolding and related health and safety requirements. Paper-based test includes questions about the scaffolding installation stages, the purpose of the vehicles used, common mistakes made during the installation stage and the use of personal protective equipment. For example, wearing hard hat, locking the diagonal pin and attaching the seat belt are common safety acts in the installation of scaffolding. Most of the questions are related to occupational safety and the correct use of materials. The aim of administering paper-based test is to measure participants' initial knowledge on scaffolding tasks and related health and safety implementations. Additionally, the participants are also asked to take the same paper-based test after completing V-SAFE v.2 training module so that the effect of the training tool on participants' learning can be measured quantitatively.

In the last part of the experiments, an evaluation questionnaire was applied to the participants. The training evaluation model developed by Kirkpatrick (1998) is taken as the source for the preparation of the questionnaire to be applied at the end of the experiment. In the survey model developed by Kirkpatrick, there are four levels of evaluation: (i) reaction, (ii) learning, (iii) behavior, (iv) results. The Kirkpatrick Model covers four main stages to determine whether a safety training is effective or not. 13 questions in the survey are based on four different levels of the evaluation in which are mainly focused on gathering the participants' points of view as suggested by Kirkpatrick (1998). The survey includes specific questions based on the participants' point of view of the training program. Survey form is presented in Appendix A. Many research areas have adopted this model in order to evaluate the effectiveness of the training program. For example, Guo *et al.* (2012) developed a game technology-based safety training on the dismantling of tower cranes and tested its validity

by adopting the interview and questionnaire survey methods. Four levels in the survey are explained in detail.

Firstly, the aim of reaction stage is to measure how participants react to V-SAFE v.2. Kirkpatrick (1998) called this step as a measure of customer satisfaction basically. It is important for the effectiveness of the training that the participants like the subject, the materials, the presentation, the area where the training is given and see this training as a special experience. In addition, the participants' thoughts on this subject and the level of perceived perception of the participants are also indicators for the quality of the training. It is also important to analyze the response to identify deficiencies in the training offered. The following 5 questions are presented to measure the reaction of the participants at the end of the experiment based on Kirkpatrick model:

- Do you agree that the dangers shown in the virtual training tool represent real situations at the construction site? (1)
- Do you think that the time you have spent during the virtual training tool was efficient? (2)
- Do you agree that the virtual safety training is an entertaining experience? (3)
- Do you think that this kind of trainings can be adopted for other training activities? (4)
- Do you think that the information you have learned in the virtual safety training is understandable and permanent? (5)

In the learning stage, what the participants learn during the training was measured. It is evaluated in which extent participants change the attitudes towards the safety issues or increase the knowledge as a result of the training provided (Kirkpatrick, 1998).
At this stage, the following questions in the survey are categorized as the learning stage;

- Do you agree that the virtual safety training has increased your knowledge on the scaffolding installation and related work safety issues? (6)
- Do you agree that you were concentrated during the virtual safety training tool? (7)
- Do you think that the virtual safety training tool experience was consistent with your experience in the scaffolding installation? (8)

- Do you agree that the training you have received is a qualified training? (9)

In the behavior stage, the participants were evaluated to see whether there is any change in the behavior and thus how the participants turn the knowledge they have learned into practice.

The following questions represent the behavior stage of Kirkpatrick model;

- Do you think that you will remember the information you have learnt from this virtual safety training tool in the future? (10)
- Do you recommend similar trainings to your friends? (11)
- Do you want to receive a similar training in the future? (12)
- Do you think the training you received will help prevent accidents at the construction site? (13)

According to Kirkpatrick (1998), result stage covers results that occurred as a result of attending the program which may include improved quality, reduced severity of accidents, and increased production. At this stage, the reaction of the participants to the training, the information they learned during the training and the continuity of the information learned, whether they apply the information they learned after the training are analyzed. The effectiveness of V-SAFE v.2 is evaluated by assessing V-SAFE scores. As a result of the evaluation, the effectiveness of the training given at different levels is observed and new training strategies can be developed according to these results. The survey results are then collected in the excel document and assessed in order to see the participants' experience in V-SAFE v.2.

3.4. Analysis Methods

Analysis methods have been determined in line with the characteristics of the data to be used for the hypotheses in Hypothesis Development section. Moreover, similar studies have been reviewed, and analysis methods used in such studies have been also considered. The characteristics of the data and the variables are the determinants in selecting the

appropriate statistical analysis methods (Murrar and Brauer, 2018). The appropriate analysis methods for each hypothesis are briefly introduced in this section.

As hypotheses are listed in the Hypothesis Development section, there are five hypotheses covering the aim and objectives of this thesis. To test the three hypotheses, the paper-based test scores of three subject groups before and after receiving V-SAFE v.2 training module are compared with appropriate statistical methods separately. In this way, the effectiveness of V-SAFE v.2 on three subject groups' learnings can be investigated. The effectiveness of a training can be measured based on the test results before and after training (e.g. Sacks *et al.*, 2013; Pedro *et al.*, 2016; Reid and Dawson, 2001). Sacks *et al.* (2013) compared the subject group's safety test scores collected just prior to the training with the scores collected after the training by using T-tests with a degree of significance 95%. They adopted a T-test comparison to confirm that there was no significant difference in the participants' initial safety knowledge and risk perception skills (Sacks *et al.*, 2013). Similarly, Pedro *et al.* (2016) assessed the participants' performances through a paper-based test and adopted a paired T-test in order to validate the significant difference between the before and after education scores of the participants. Reid and Dawson (2001) performed a paired T-tests of the subject groups to evaluate if there was a significant difference in performance between the shifts for both the older and younger subject groups in their study. Considering the relevant studies, a paired T-test is commonly adopted in similar cases (e.g. Sacks *et al.*, 2013; Pedro *et al.*, 2016; Reid and Dawson, 2001). Considering the similar studies in the literature, a paired T-test is determined to test if there is a significant difference on the increase of the participants' knowledge after receiving V-SAFE v.2.

Depending on the characteristics of the data (parametric or nonparametric), three tests are considered: (i) the paired *t*-test, (ii) Wilcoxon signed-rank test and (iii) the sign test (Dytham, 2011). Hypotheses 1,2 and 3 evaluate the effectiveness of V-SAFE v.2 on the subject group's knowledge, however, the subject group differs in each hypothesis. Therefore, the assumptions for paired T-test should be checked for the data of each hypothesis before applying the analysis. The analysis methods are applied in SPSS. For Hypothesis 1, the data meets the assumptions for paired T-test. However, since the assumptions for paired T-test are not met for the data set in Hypothesis 2 and 3, Wilcoxon signed-rank test, which is the non-parametric equivalent of paired T-test (Dytham, 2011), is adopted.

Hypothesis 4 and 5 test the appropriateness of using students as surrogates for construction workers and engineers in the development of the virtual reality safety training tool, respectively. Use of students as surrogates for construction workers and engineers can be evaluated by the structured experimental setting with appropriate analysis methods. Using students as surrogates for construction workers has been adopted in studies related to development of the safety training tool (e.g. Guo *et al.*, 2012 and Perlman *et al.*, 2014). For example, Guo *et al.* (2012) used a technology-based safety training in the experiments with 15 managers so that managers can capture the potential hazard situations by themselves in the virtual environment. Perlman *et al.* (2014) used two-way ANOVA to compare the two test methods (traditional test and virtual test) over the subject groups. In this regard, the experiments in this study are designed to be held with the three subject groups: senior civil engineering students, engineers and construction workers by grouping them as students and construction workers in Hypothesis 4 and students and engineers in Hypothesis 5. Additionally, two-factor mixed ANOVA is used for Hypothesis 4 and Hypothesis 5.

3.4.1. Paired T-test

Paired T-test is used when there are two observations and same participants take place in both conditions of the experiment (Walpole *et al.*, 2012). Paired comparisons are used for the data sets in which a single individual is tested twice (i.e. before and after) or a sampling data re-tested (Dytham, 2011). The data set in a paired t-test are the differences between pairs (Elliot and Woodward, 2007). In a paired T-test $\mu 1$ and $\mu 2$ stand for the population means of the before and after observations of a subject group (Elliott and Woodward, 2007). For a paired T-test, the difference between the means of two data sets (i.e., $\mu 1 - \mu 2$) equal to the population mean of the difference scores, denoted μd (Elliott and Woodward, 2007). The aim of paired t-test is to assess if there is a significant difference between two group means which has been collected by means of a related measures design (Kerr *et al.*, 2002). Therefore, the null and alternative hypotheses are stated as (Elliott and Woodward, 2007):

H_0: $\mu d = 0$ (the population mean of the differences is zero).
H_1: $\mu d \neq 0$ (the population mean of the differences is not zero).

In Hypothesis 1, since before-and-after observations are on the same subjects (i.e. students' paper-based test results before and after receiving V-SAFE v.2 training module), the samples are not independent. The two populations in this case are considered as "before" and "after," and the experimental unit is considered as the individual (Walpole *et al.*, 2012). For Hypothesis 1, to test the null hypothesis that the mean difference is zero, paired T-test is used in SPSS statistics package. The difference between the two observations on each experimental unit (the individual) is calculated and the mean differences and the standard deviation of the differences are calculated accordingly (Walpole *et al.*, 2012). It should be noted that, this test can be valid if the differences are normally distributed (Walpole *et al.*, 2012). Meaning that, experimental units should be relatively homogeneous, and each unit should experience both population conditions which in turn reduces the effective experimental error variance (Walpole *et al.*, 2012).

The *t*-test is a parametric test and it makes the following assumptions (Kerr *et al.*, 2002):

- The level of measurement of the dependent variable must be at least interval.
- The dependent variable is normally distributed.
- The variances of the samples are not significantly different.

For Hypothesis 1, since the data meet all assumptions for a paired T-test, the paired T-test is adopted by using SPSS statistics package. Assumptions and the results are made by interpreting the tables and the values generated in SPSS statistics package properly.

3.4.2. Wilcoxon Signed-Rank Test

For Hypotheses 2 and 3, although the concept is quite similar with Hypothesis 1, the paired T-test cannot be applied since the assumptions are violated in these two hypotheses because the test scores for the engineer and worker data are found to be non-parametric. Therefore, non-parametric equivalent of paired T-test is appropriate to assess these hypotheses. Wilcoxon signed-rank test corresponds to the non-parametric test of the paired *t*-test (Elliott and Woodward, 2007 and Dytham, 2011). This test requires fewer assumptions about the shape of the data, however, it assumes that the data are continuous (Dytham, 2011) and quantitative (Elliott and Woodward, 2007). In addition, at least six pairs of data are

required before the test can be carried out (Dytham, 2011). Hypothesis 2 and 3 meet the assumptions for Wilcoxon signed-rank test. The absolute values of the differences are used in the Wilcoxon signed-rank test (Elliott and Woodward, 2007). This test ranks the absolute differences and the sum of the ranks for positive differences is compared with the sum of the ranks for the negative differences (Elliott and Woodward, 2007). Therefore, the null and alternative hypotheses are stated as (Elliott and Woodward, 2007):

H_0: The probability of a positive difference is equal to the probability of a negative difference.
H_1: The probability of a positive difference is not equal to the probability of a negative difference.

3.4.3. Two-factor Mixed ANOVA

Two-factor ANOVA is a powerful test and appropriate to be used for the cases in which there are two independent ways of making the observations into groups and there is more than one observation per factor combination (Dytham, 2011). Any feature of the experiment is denoted as the term “factor” and the actual values used are defined as the levels of a factor (Walpole *et al*, 2012). The two-factor ANOVA consists of the main effects for each factor and an interaction effect (Elliott and Woodward, 2007). Interaction effect measures if two or more grouping variables have an interaction (additive) effect or not (Dytham, 2011).

There are three tests in two-factor mixed ANOVA: (i) test for interaction effect; (ii) two tests for main effects (Elliott and Woodward, 2007). These three tests are adopted in an ANOVA table as F-tests. A low p-value (usually less than 0.05) for a test stands for the evidence to reject the null hypothesis and accept the alternative hypothesis (Elliott and Woodward, 2007).

A mixed model analysis of variance (or mixed model ANOVA) is used for the experimental data in which includes (i) a continuous dependent variable, (ii) two or more categorical independent variables, (iii) at least one independent variable as between-subjects

factor, and (d) at least one independent variable as within-subjects factor (Murrar and Brauer, 2018).

There are several assumptions that needs to be checked before applying two-factor ANOVA: (i) the data set must be continuous, (ii) the data set must be at least approximately normally distributed and (iii) the variation is the same (Dytham, 2011). Before applying two-factor ANOVA, the data should meet the assumptions of two-factor mixed ANOVA. Even if the data does not meet the normality distribution, the analysis methods for parametric distributions can be still applied for the large sample sizes (>30 or 40) (Elliott and Woodward, 2007). The concepts of skewness and kurtosis are used as statistical means of assessing the normality assumption (Kerr *et al.*, 2002). In this case, skewness and kurtosis are two main ways to look at the deviation from the normality for the data (Ghasemi and Zahediasl, 2012).

Skewness is used as a measure of the symmetry of the distribution and it is only used for the data set where there are more than 30 observations (Dytham, 2011). A zero value of skewness means that the data set is symmetrical and If the data set is symmetrical then the value of skewness will be 0 (Dytham, 2011). However, positive value of skewness means that there is a tail to the right and negative value of skewness means that there is a tail to the left (Dytham, 2011).

Kurtosis is used as a measure of the 'flatness' of the data (Dytham, 2011). A symmetrical distribution can differ from the normal in being either leptokurtic or platykurtic in which if there are more observations closer to the mean and in the tails, the distribution is called leptokurtic and if there fewer observations closer to the mean and tails, the distribution is called platykurtic distribution (Dytham, 2011).

In Hypothesis 4 and Hypothesis 5, it is logical to look at a value of skewness and kurtosis values since there are 40 observations in both pre-test and post test data. If the kurtosis and skewness values are divided by their standard errors (SE_{skew} and $SE_{kurtosis}$) and the calculated z-value ($Z_{kurtosis}$ and $Z_{skewness}$) are not greater than ±1.96, the data does not significantly deviate from a normal distribution and the distribution is normal with at least 95% confidence (Kerr *et al.*, 2002).

$$Z_{skew} = Skew / SE_{skew} \tag{3.1}$$

$$Z_{kurtosis} = Kurtosis / SE_{kurtosis} \tag{3.2}$$

These calculations are made before applying two-factor mixed ANOVA for Hypothesis 4 and 5.

4. RESULTS

The data set collected from 60 participants consists of 20 civil engineering students, 20 construction workers and 20 engineers. Five hypotheses are tested with the appropriate statistical analysis methods and results are stated for each hypothesis accordingly. The statistical analysis methods are adopted after checking corresponding assumptions in SPSS. For Hypothesis 1, 2 and 3, the SPSS outputs of normality tests are presented in Appendix B. Necessary statistical checks for normality assumptions in two-factor mixed ANOVA for Hypothesis 4 and 5 are presented in Appendix C.

For Hypothesis 1, the effectiveness of V-SAFE v.2 on the student subject group's learnings is investigated by comparing the paper-based test results before and after receiving V-SAFE v.2. The data set consists of test scores of 20 civil engineering students before and after receiving V-SAFE v.2 training module. Paper-based test scores collected prior to the training module are named as "Pre_Test_Score" and paper-based test scores collected after the training module are named as "Post_Test_Score" in SPSS. In this case, paired T-test is adopted to test the null hypothesis of Hypothesis 1.

Assumption checks for paired T-test are as the following: The paper-based test scores are assigned as dependent variable and it is continuous. Meaning that, the first assumption for paired T-test is satisfied. For the normality assumption of paired T-test, since there are 20 observations in this case, the normality checks are made by using the Shapiro-Wilk tests instead of the skewness and kurtosis values. The pre-test scores and post-test scores of students are found as normally distributed according to the Shapiro-Wilk test results, since the p value is greater than 0.05 (p = 0.162 for pre-test scores and p=0.147 for post-test scores). Therefore, the normality assumption of paired T-test is met for the data set in Hypothesis 1. Since all assumptions for the paired T-test are met for this data set, paired T-test can be applied to test the null hypothesis in Hypothesis 1.

Two variables are assigned as "Pre_Test_Score" and "Post_Test_Score" in SPSS. Mean, number of observations (N) and standard deviations for the pre-test scores and post-

test scores are generated from SPSS in Table 4.1. Mean of pre-test scores are 5.80 while mean of post-test scores is 8.05.

Table 4.1. Paired Samples Statistics for Paper-based Test Scores of Students.

Pair 1	**Mean**	**N**	**Std. Deviation**
Pre_Test_Score	5.80	20	1.436
Post_Test_Score	8.05	20	1.146

In Table 4.2, statistics for paired difference are presented as mean, standard deviation the paired T-test result, degree of freedom and the p-value, respectively. The mean difference between pairs of data (pre-test score and post-test score) are reported as -2.25. The t-test statistic is -7.336. In the last column, the sig. value (*p* value) approximates 0 which is smaller than 0.05 and therefore the null hypothesis is rejected. Meaning that, there is a significant difference between the pre-test scores and post-test scores for the student subject group.

Table 4.2. Paired Samples Test for Paper-based Test Scores of Students

.

Pair 1	**Paired Difference**				
	Mean	**Std. Deviation**	**t**	**df**	**Sig. (t-tailed)**
Pre_Test_Score - Post_Test_Score	-2.25	1.371	-7.336	19	0.000

For Hypothesis 2, paired t-test is not adopted because the data of post-test scores are found to be not normally distributed according to Shapiro-Wilk test of normality which has a p value less than 0.05 (p = 0.0496). As a result, since the data is not normal, Wilcoxon signed-rank which is the nonparametric equivalent for paired t-test is adopted to test the Hypothesis 2. For Hypothesis 2, the dependent variable is construction workers' paper-based test scores. The pre-test and post-test scores for construction workers are assigned as "Pre_Test_Score" and "Post_Test_Score" in SPSS. Before applying Wilcoxon signed-rank test, its assumptions should be checked appropriately.

Assumption checks for Wilcoxon signed-rank test are as the following: To be able to apply Wilcoxon signed-rank test, the assumptions that presented in the previous chapter are checked. In this case, the first assumption is met since each pair represents each participant which is independent from each other. The dependent variable is the paper-based test scores and continuous; therefore, the second assumption is met. Lastly, there are 20 paired test scores for 20 participants. As a result, since all assumptions are met, Wilcoxon signed-rank test is used to test Hypothesis 2.

Descriptive statistics for the data summarized the number of observations, mean, standard deviations for pre-test and post-test scores of construction workers in Table 4.3. For construction workers, means of the pre-test and post-test scores are found as 6.35 and 7.85, respectively.

Table 4.3. Descriptive Statistics for Paper-based Test Scores of Construction Workers.

Variables	**N**	**Mean**	**Std. Deviation**
Pre_Test_Score	20	6.35	1.631
Post_Test_Score	20	7.85	1.387

In Table 4.4, Wilcoxon signed-rank test categorizes the paired data into three categories. Negative Ranks represents that post-test score is less than pre-test scores and there is no case that post-test scores are lower than pre-test scores. Positive Ranks represents post-test score is greater than pre-test scores and 16 cases are found that post-test scores are higher than pre-test scores. Ties represents that two scores are equal. In this case, 4 cases are found that post-test scores are equal to pre-test scores.

Table 4.4. Wilcoxon Signed-Rank Test for Hypothesis 2.

	Ranks	**N**	**Mean Rank**	**Sum of Ranks**
Post_Test_Score - Pre_Test_Score	Negative Ranks	0	0.00	0.00
	Positive Ranks	16	8.50	136.00
	Ties	4	-	-
	Total	20	-	-

In Table 4.5, Z-value is presented coupled with a “Asymp. Sig. (2-tailed)” which represents the p-value (Dytham, 2011). In this case, since the p-value approximates to zero which is less than 0.05, so the null hypothesis must be rejected. The alternative hypothesis that the pre-test and post-test results of construction workers are significantly different is accepted.

Table 4.5. Test Statistics for Paper-based Test Scores of Construction Workers.

Test Statistics	**Post_Test_Score - Pre_Test_Score**
Z	-3.581
Asymp. Sig. (2-tailed)	0.000

For Hypothesis 3, paired t-test is not adopted because, Shapiro-Wilk normality test of normality reveals that the result is significant since the with p-values of both pre-test scores and post-test scores are smaller than 0.05 (p= 0.19 for pre-test scores and p=0.05 for post-test scores). Therefore, Wilcoxon signed-rank test, which is the nonparametric equivalent for paired t-test, is be adopted to test the Hypothesis 3. In this case, the dependent variable is engineers' paper-based test scores. The data collection for pre-test scores and post-test scores are assigned in SPSS as “Pre_Test_Score” and “Post_Test_Score”, respectively.

Assumption checks for Wilcoxon signed-rank test are as the following: In this case, each pair is independent (a pair for each worker) which meets the first assumption, the dependent variable is continuous and there should be more than 5 paired data. As a result, three assumptions are met, therefore Wilcoxon signed-rank test is used to test Hypothesis 3.

Descriptive statistics are presented in Table 4.6. Mean of pre-test and post-test scores for engineers are found as 7.35 and 8.90, respectively.

Table 4.6. Descriptive Statistics for Paper-based Test Scores of Engineers.

Variable	**N**	**Mean**	**Std. Deviation**
Pre_Test_Score	20	7.35	0.988
Post_Test_Score	20	8.90	0.788

In Table 4.7, Wilcoxon signed-ranks test results are presented. In this case, there is no case that post-test scores are lower than pre-test scores. 18 cases are found that post-test scores are higher than pre-test scores. Lastly, 2 cases are found that post-test scores are equal to pre-test scores.

Table 4.7. Wilcoxon Signed-Rank Test for Hypothesis 3.

	Ranks	**N**	**Mean Rank**	**Sum of Ranks**
Post_Test_Score - Pre_Test_Score	Negative Ranks	0	0	0.00
	Positive Ranks	18	9.50	171.00
	Ties	2	-	-
	Total	20	-	-

In Table 4.8, Z-value is presented coupled with a ‘Asymp. Sig. (2-tailed) which represents the p-value (Dytham, 2011). In this case, since $p < 0.05$, so the null hypothesis should be rejected. The alternative hypothesis that the pre-test and post-test results of engineers are significantly different is accepted.

Table 4.8. Test Statistics for Paper-based Test Scores of Engineers.

Test Statistics	**Post_Test_Score - Pre_Test_Score**
Z	-3.794
Asymp. Sig. (2-tailed)	0.000

For Hypothesis 4, two-factor mixed ANOVA is used to test Hypothesis 4 because there is one between subject factor and one within subject factor. This statistical analysis is made in SPSS. The within-subjects factor is assigned as “Performance” and it has two levels, pre-test scores and post-test scores. Pre-test score data consists of the number of correct answers of students and of construction workers in the paper-based test before receiving V-SAFE v2. Similarly, between-subject factor is assigned as “Subject_Type” in SPSS and this independent variable has also two levels: 1 is for students and 2 is for workers.

Post-test score consists of the number of correct answers of students and of construction workers in the paper-based test after receiving V-SAFE v2. There are 20 students and 20 construction workers who attended the experimental settings; therefore, the size of the data is 40 for both pre-test score and post-test score data. Since there are fewer than three groups, post-hoc tests are not performed for subject type.

Assumption checks for Two-factor mixed ANOVA are as the following: Shapiro-Wilk test of normality is generated the significance values for both pre-test score and post-test score data as smaller than 0.05. Meaning that, the data is not normally distributed. However, even if the data does not meet the assumption of normality distribution, the analysis methods for parametric distributions can be still applied for the large sample sizes (>30 or 40) (Elliott and Woodward, 2007). Therefore, In Hypothesis 4, it is logical to look at a value of skewness to determine the symmetry since there are 40 observations in both pre-test and post test data. Considering these ranges for Skewness, skewness is 0.221 for pre-test scores and 0.017 for post-test scores. The Skewness value and the standard error of Skewness (SE_{skew}) are given for pre-test and post-test scores generated in SPSS. Putting the skewness value and SE_{skew} in the formula (3.1), $Z_{skewness}$ are found as 0.59 and 0.045 for pre-test score data and post-test score data, respectively. Both values are in a specified range of ±1.96 with 95% confidence. It means that the data for pre-test scores does not significantly deviates from normality.

To check the kurtosis values, the Kurtosis values and the standard error values of Kurtosis for pre-test and post-test scores generated in SPSS are used in the formula (3.2) to find the values of $Z_{kurtosis}$ for pre-test and post-test scores. $Z_{kurtosis}$ is found as -0.556 and -1.296 for pre-test score data and post-test score data, respectively. Both values are in a specified range of ±1.96 with 95% confidence. It means that the data for pre-test scores does not significantly deviates from normality as well.

To check the homogeneity of variance assumption for the data, Levene's Statistic is used for mixed ANOVA which should be assessed at all levels of the within-subject variable (both pre-test score and post-test score). If the p-value for this data is significant, the assumption is not met. In this case, p-values generated in SPSS is greater than 0.05 which is non-significant result, therefore the homogeneity of variance assumption is met.

Mauchly's Test of Sphericity is generated for the assumption of the sphericity which is used where within-subject variable has more than two levels. Therefore, there is no need to look at this table, either.

For Hypothesis 4, the information of pre-test scores and post-test scores are given in Table 4.9. The mean, standard deviation and the size of the data are summarized in Table 4.9. The mean scores for pre-test results are 5.80 and 6.35 for students and workers, respectively. The mean scores for post-test results are 8.05 and 7.85 for students and workers, respectively.

Table 4.9. Descriptive Statistics of Paper-based Test Scores in Hypothesis 4.

Variable	Subject_Type	Mean	Std. Deviation	N
Pre_Test_Score	Students	5.80	1.436	20
	Workers	6.35	1.631	20
	Total	6.08	1.542	40
Post_Test_Score	Students	8.05	1.146	20
	Workers	7.85	1.387	20
	Total	7.95	1.260	40

In SPSS, the Box’s Test of Equality of Covariance Matrices and Multivariate Tests are generated. However, since there is no multiple dependent variable, there is no need to interpret these tests.

Table 4.10 shows the ANOVA results for the within-subjects variable (performance) and any interactions with that variable. The values in the Sphericity Assumed rows are considered only. The second column of Table 4.10 called df represents the degrees of freedom. In this case the degrees of freedoms are shown in the Sphericity assumed rows. In the fourth column of Table 4.10, F stands for F-ratio statistics calculated by the ANOVA. The last column in Table 4.10, Sig stands for Significance Level which is used to determine if the null hypothesis is true. The p-value should be less than 0.05 for the F-ratio to be significant. In other words, if p-value is less than 0.05, the null hypothesis is rejected, and it is concluded that the independent variables has a significant effect. In Table 4.10, the

p-value for within-subject variable, performance, is less than 0.05. Meaning that, performance has a significant effect. On the other hand, the p-value for the interaction effect is greater than 0.05, so the interaction term is not statistically significant.

Table 4.10. Test Results of Performance Effect in Hypothesis 4.

	Sphericity Assumed				
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Performance	70.312	1	70.312	97.603	0.000
Performance * Subject_Type	2.813	1	2.813	3.904	0.055
Error(Performance)	27.375	38	0.720	-	-

Table 4.11 gives the ANOVA statistics for the between-participants variable (subject type). There is no significant main effect of subject type on the test results since the sig value is greater than 0.05 (p=0.667) as shown in Table 4.11.

Table 4.11. Test Results of Subject Type Effect in Hypothesis 4.

Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Intercept	3934.012	1	3934.012	1206.801	0.000
Subject_Type	0.612	1	0.612	0.188	0.667
Error	123.875	38	123.875	-	-

For Hypothesis 5, the analysis is pretty much same with the Hypothesis 4, however, this time the subject groups are students and engineers instead of the construction workers. Therefore, two factor mixed ANOVA is used for this analysis in SPSS statistics package as well because there is one between subject factor and one within subject factor. Subject type is assigned as between subject factor and performance is assigned as within subject factor. Both factors have two levels: students and engineers for subject type and pre-test score and

post test score for performance, respectively. Since there are fewer than three groups which is only two groups, post-hoc tests are not performed for Subject type. There are 20 students and 20 engineers who involved in the experimental settings; therefore, the size of the data is 40 for both pre-test score and post-test score data with no missing values.

Assumption checks for Two-factor mixed ANOVA are as the following: The p-values for pre-test score and post-test score data as a result of normality tests are generated in SPSS. The p-value of Shapiro-Wilk test for pre-test scores is greater than 0.05, meaning that the data is normally distributed for this test. However, the p-values of Shapiro-Wilk tests for post-test score is smaller than 0.05. Meaning that, the post-test score data for the two subject groups is not normally distributed according to the tests of Normality. Since both the pre-test score and post-test score data have 40 observations, two factor mixed ANOVA can be adopted to test Hypothesis 5 if the data does not significantly deviate from the normality. However, since the post-test data does not meet the normality for Shapiro-Wilk test, the deviation from the normality should be checked for the data.

Considering the range for skewness and kurtosis values ($Z_{skewness}$ and $Z_{kurtosis}$) for the normality, skewness is -0.210 for pre-test scores and -0.540 for post-test scores. Using the formula (3.1), the Skewness values and the standard error values of Skewness for pre-test and post-test scores are generated in SPSS, the value of Z_{skew} is found as -0.56 for pre-test score data and -1.44 for post-test score data. Both values are in a specified range of ±1.96 with 95% confidence. It means that the data for pre-test scores does not significantly deviates from normality.

Kurtosis values are presented as -0.178 for pre-test score and -0.118 for post-test score. Kurtosis values and the standard error values of Kurtosis are put in the formula (3.2) to find the value of $Z_{kurtosis}$. The values of $Z_{kurtosis}$ are found as -0.24 and -0.16 for pre-test score data and post-test score data, respectively. Both values are in a specified range of ±1.96 with 95% confidence. It means that the data for pre-test scores and post-test scores does not significantly deviates from normality as well.

The Levene's Statistic is used to check the homogeneity of variance assumption for the data. This should be applied at all levels of the within-subject variable (both pre-test and

score and post-test score) for mixed ANOVA. In this data, the p-value is greater than 0.05 which is non-significant result, therefore the homogeneity of variance assumption is met for both pre-test score and post-test score.

Table 4.12 includes the mean, standard deviations and the size of the data which gives information about the pattern of the pre-test and post-test scores for students and engineers. The mean scores for pre-test results are 5.80 and 7.35 for students and engineers, respectively. The mean scores for post-test results are 8.05 and 8.90 for students and engineers, respectively.

Table 4.12. Descriptive Statistics of Paper-based Test Scores in Hypothesis 5.

Variable	**Subject_Type**	**Mean**	**Std. Deviation**	**N**
Pre_Test_Score	Students	5.80	1.436	20
	Engineers	7.35	0.988	20
	Total	6.58	1.448	40
Post_Test_Score	Students	8.05	1.146	20
	Engineers	8.90	0.788	20
	Total	8.48	1.062	40

Table 4.13 shows the ANOVA results for the within-subjects variable (performance) and any interactions with that variable. As it is explained in Hypothesis 4, the values in the sphericity assumed rows are considered only. The last column in Table 4.13, Sig stands for significance level which is used to determine if the null hypothesis is true. The p-value should be less than 0.05 for the F-ratio to be significant. Therefore, if p-value is less than 0.05, it is concluded that the independent variables has a significant effect.

In Table 4.13, the p-value for within-subject variable, performance, is less than 0.05. Meaning that, performance has a significant effect. On the other hand, the p-value for the interaction effect is greater than 0.05, so the interaction term is not statistically significant.

Table 4.13. Test Results of Performance Effect in Hypothesis 5.

	Sphericity Assumed				
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Performance	72.200	1	72.200	100.314	0.000
Performance * Subject_Type	2.450	1	2.450	3.404	0.073
Error(Performance)	27.350	38	0.720	-	-

Table 4.14 summarized the ANOVA statistics for the between-participants variable (subject type). There is a significant main effect of subject type on the test results since the sig value is smaller than 0.05 in Table 4.14.

Table 4.14. Test Results of Subject Type effect in Hypothesis 5.

Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Intercept	4530.05	1	4530.05	2563.54	0.000
Subject_Type	28.80	1	28.80	16.30	0.000
Error	67.15	38	1.77	-	-

5. DISCUSSION

Five hypotheses are determined based on the aim and objectives of this thesis. The first objective is to analyze the effectiveness of V-SAFE v.2 on three subject groups' learnings regarding the topics covered in V-SAFE v.2 training module. The subject groups consist of civil engineering students, construction workers and engineers. The effectiveness of V-SAFE v.2 is investigated for each subject group which constitutes three hypotheses accordingly. Results regarding the effectiveness of the V-SAFE v.2 are discussed in the first section. The appropriateness of using students as surrogates is investigated for construction worker subject group and engineer subject group in Hypothesis 4 and 5, respectively and results are discussed in the second section. In the third section, V-SAFE v.2 performances of three subject groups coupled with their demographic information and feedbacks regarding V-SAFE v.2 are represented graphically to make general comments as supplementary findings. Limitations and future directions for this thesis are briefly explained in the last section.

5.1. Effectiveness of V-SAFE v.2

Insufficient training methods are major contributors to high rates of accidents in construction sector (Wilkins, 2011; *Guo et al.*, 2012). Many researchers questioned the effectiveness of current training tools provided construction workers in their studies (e.g. Wilkins, 2011; Sacks *et al.*, 2013; Mohd *et al.*, 2019). Effective training methods should be adopted in safety training tools for construction workers to enhance the construction safety. Virtual environments provide its users with a virtual construction site including representative risks of real-life and experience and the consequences of an unsafe act without any danger (Pedro *et al.*, 2016). Although many studies focused on the effectiveness of safety training tools developed in virtual environments, there is a limited empirical finding to prove the effectiveness of computer aided technologies in the accident rate reduction (Gao et al., 2019).

Hypothesis 1, 2 and 3 investigate the effectiveness of V-SAFE v.2 on three subject groups' learning on the topics covered in the training module of V-SAFE v.2: student subject

group for Hypothesis 1, construction worker subject group for Hypothesis 2, and engineer subject group for Hypothesis 3. Experimental flow and the materials used in the experiments remain same for all hypotheses, however, the only difference is their subject groups and analysis methods depending on the data characteristics of each hypothesis. The data used in the first three hypotheses are the pre-test scores and post-test scores of the subject groups by collecting the paper-based test results. Paper-based test consists of 10 questions covering the basic tasks in scaffolding installation and related health and safety implementations which are presented in V-SAFE v.2 training module. Pre-test scores represents the number of correct answers in the paper-based test before receiving V-SAFE v.2 training tool. The aim of pre-test is to measure the initial knowledge of students on the subjects covered in V-SAFE v.2 training tool. The number of correct answers in the paper-based test after receiving V-SAFE v.2 training tool are assigned as post-test scores. Pre-test scores and post-test scores are analyzed in SPSS by adopting appropriate statistical methods and the results are interpreted for each hypothesis accordingly.

For Hypothesis 1, 20 civil engineering students at Boğaziçi University attended the experiments. The test scores for the same subject groups that are collected at two different times can be tested by paired T-test (Walpole *et al.*, 2012). Therefore, paired T-test is adopted to test the null hypothesis for Hypothesis 1 after checking the assumptions for paired T-test. The mean of pre-test scores for civil engineering students is found as 5.80, while the mean of post-test scores is found as 8.05. Considering both pre-test and post-test scores, there is a 38.79% increase in their performance on the paper-based test results after receiving V-SAFE v.2 training module. According to paired T-test results, the difference between pre-test and post-test scores for students ($t = -7.336$, $p < 0.05$) is found as significant. Meaning that, V-SAFE v.2 training module has a significant positive effect on students' learning regarding the basic scaffolding knowledge and related health and safety issues.

For Hypothesis 2, the effectiveness of V-SAFE v.2 on construction worker subject group's learning is investigated. The experiments are conducted with 20 construction workers from two construction sites. In this case, the non-parametric equivalent of paired T-test which is Wilcoxon signed-rank test (Elliott and Woodward, 2007; Dytham, 2011) is used to test the null hypothesis for Hypothesis 2 since the data is not normally distributed. The mean of pre-test scores for construction workers is found as 6.35. The mean of post-test

scores is found as 7.85. Considering the mean for both pre-test and post-test scores of construction workers, there is a 23.62% increase in their performance on the paper-based test results after receiving V-SAFE v.2 training module. According to Wilcoxon signed-rank test, the difference between pre-test and post-test scores for construction workers ($Z = -3.581$, $p < 0.05$) is found as significant. Meaning that, V-SAFE v.2 training module has a significant positive effect on construction workers' learning on basic scaffolding knowledge and related health and safety issues given in V-SAFE v.2.

For Hypothesis 3, 20 engineers from two construction companies attended the experiments. Wilcoxon signed-rank test is adopted to test the null hypothesis for Hypothesis 3 since the pre-test and post-test results of engineers are not normally distributed. The mean of pre-test scores for engineers is found as 7.35 while the mean of post-test scores is found as 8.90. Considering the means for both pre-test and post-test scores of engineers, there is a 21.09% increase in their performance on the paper-based test results after receiving V-SAFE v.2 training module. According to Wilcoxon signed-rank test, the difference between pre-test and post-test scores for construction workers ($Z = -3.794$, $p < 0.05$) is found as significant. In other words, V-SAFE v.2 training module has a significant positive effect on engineers' learning on basic scaffolding knowledge and related health and safety issues.

In Figure 5.1, the boxplot is generated to illustrate the difference between the pre-test and post-test scores for students, construction workers and engineers. The change in their safety knowledge after receiving V-SAFE v.2 can also be seen visually in Figure 5.1 below.

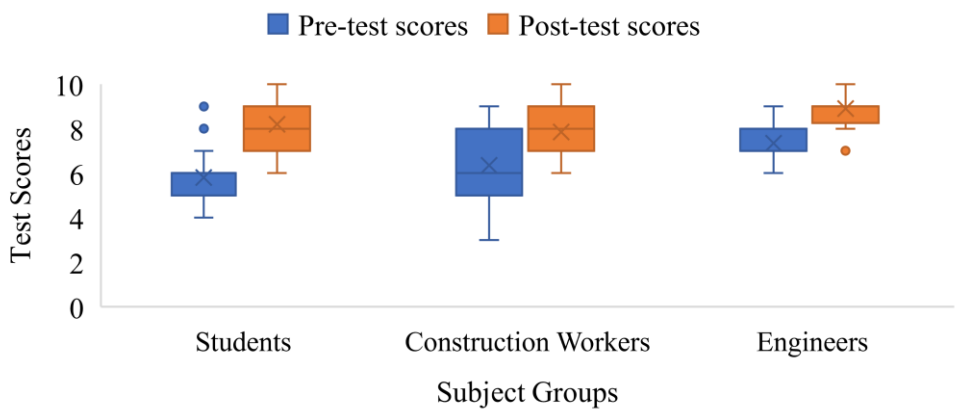

Figure 5.1. Boxplots for Pre-test and Post-test Scores of Three Subject Groups.

Considering the results for the first three hypotheses, the developed V-SAFE v.2 is found as an effective virtual safety training tool on the learnings of students, construction workers and engineers regarding the basic scaffolding tasks and related health and safety implementations covered in V-SAFE v.2 training module. Recent studies showed the similar findings (e.g. Sacks *et al.*, 2013; Pedro *et al*., 2016). For example, Sacks *et al.* (2013) analyzed the safety test scores of construction workers and junior civil engineering students. They compared the pre-training and post-training scores of two groups and proved the effectiveness of virtual reality safety training tool on both groups' hazard identification skills (Sacks *et al.*, 2013). Additionally, Pedro *et al.* (2016) evaluated the students' performances on the developed virtual construction safety education system by comparing paper-based test results of students before and after training. They concluded that the developed virtual construction safety education system has a positive impact on improving students' safety knowledge, and hazard identification abilities (Pedro *et al.*, 2016). As a result, the effectiveness of virtual safety training tools on increasing safety knowledge and skills have been validated in many research areas as well.

5.2. Use of Students as Surrogates for Construction Workers

The appropriateness of using students as surrogates in the development of virtual safety training tools is investigated in Hypothesis 4 and Hypothesis 5. Both hypotheses investigate the use of students as surrogates, however, the subject group to be surrogated is different in the hypotheses. In hypothesis 4, the use of students as surrogates is investigated for construction workers. On the other hand, hypothesis 5 investigates the use of students as surrogates for engineers. Same methodology is applied in both hypotheses. For both hypothesis, test scores of the subject groups in paper-based tests are considered by grouping as pre-test scores and post-test scores. Two-factor mixed ANOVA is adopted to test the null hypothesis for hypothesis 4 and 5. There is one between and one within factor in this study which fit the conditions for two-factor mixed ANOVA. The within-subjects factor is assigned as "Performance" which represents the pre-test and post-test scores of the subject groups. Therefore, "Performance" has two levels: pre-test scores and post-test scores. The between-subject factor is assigned as "Subject_Type" in SPSS statistical program.

For Hypothesis 4, the subject groups are civil engineering students and construction workers. The aim is to investigate if students are appropriate subjects for construction workers in the development of virtual reality safety training tools. This independent variable which is subject type has also two levels: civil engineering students and construction workers. According to the ANOVA analysis, within-subject variable called “Performance” has a significant effect on the paper-based test results ($F = 97.603$; $p < 0.05$). Meaning that, the pre-test and post-test results of two subject groups are significantly different which proves the effectiveness of V-SAFE v.2. On the other hand, for the between-subjects variable, two-factor mixed ANOVA analysis reveals that there is no significant main effect of subject type on the test results ($F = 0.188$; $p > 0.05$). In other words, safety performances are not significantly different between two subject groups. As a result, students and construction workers showed similar safety performances on the paper-based tests.

For Hypothesis 5, the safety performances of civil engineering students are compared with the safety performances of engineers by analyzing their paper-based test results. In this case, the subject groups consist of students and engineers. The aim is to investigate if students are appropriate subjects for engineers in the development of virtual reality safety training tools. The between variable assigned as subject type has two levels: civil engineering students and engineers. Two-factor mixed ANOVA is applied to test the null hypothesis for hypothesis 5. According to two-factor mixed ANOVA results, within-subject variable, “Performance” has a significant effect on the paper-based test results ($F=100.314$; $p < 0.05$). Therefore, the pre-test and post-test results of two subject groups are significantly different which reveals the effectiveness of V-SAFE v.2. On the other hand, for the between-subjects variable, which is subject type, two-factor mixed ANOVA analysis reveals that there is a significant main effect of subject type on the test results ($F=16.298$; $p < 0.05$). Therefore, safety performances are significantly different between two subject groups. Meaning that, students and engineers did not perform similar results on the paper-based tests. So, students are not appropriate surrogates for engineers in the development of virtual safety training tools.

Consequently, combining two results for Hypothesis 4 and 5, students and construction workers performed similarly in the paper-based tests. However, students did not perform similar with engineers in the paper-based results. Safety performances for

students, construction workers and engineers in pre-test and post-test scores are summarized in Figure 5.2.

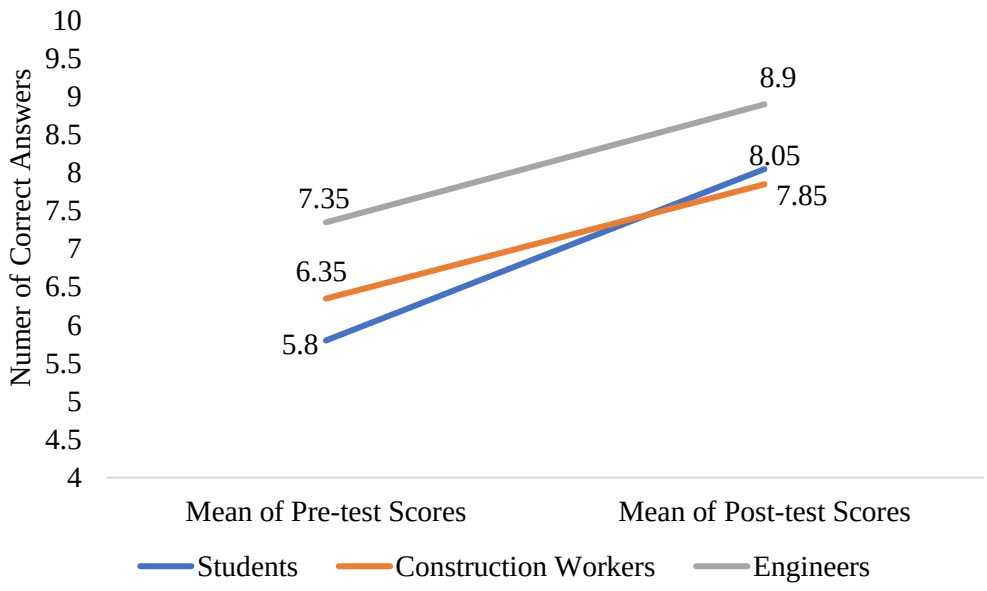

Figure 5.2. Safety Performances of Three Subject Groups on Paper-based Tests.

Considering three subject groups' safety performances on the paper-based tests, students can be used as surrogates for construction workers in the development of virtual safety training tools. However, students are not appropriate surrogates for engineers in the development of virtual safety training tools. In terms of three subject groups' learnings, the slopes of the lines presented in Figure 5.2 have been compared. For construction worker and engineer subject groups, the slopes are found as 1.5x and 1.55x respectively, while it is found as 2.25x for student subject group. Meaning that, the steepest line belongs to the student subject group. In this regard, the learning capacity of students is found as better comparing with the construction workers and engineers. Lokka and Çöltekin (2019) has also studied on the effect of age on learnings and spatial knowledge acquisition in virtual environments by using younger and older participants in their experiments. In accordance with their results, younger participants performed more successfully in virtual environments (Lokka and Çöltekin, 2019). Since the student subject group consists of the youngest participants comparing with the engineer and construction worker subject groups, the trends in Figure 5.2 are in line with the study of Lokka and Çöltekin (2019). In addition to this, since their

performances have been tracked via paper-based tests, student subject group has more familiar with this multiple-choice question format, therefore they might interpret some of the correct answers by eliminating other choices. As a result, their learning trend ended up with the steepest line comparing with construction workers and engineers.

Students are widely used as surrogates for construction workers in the development of virtual training tools in the construction research areas (e.g. Becker and Morawetz, 2004; Irizarry and Abraham, 2005; Shi *et al.*, 2018) although the appropriateness of this practice has not been widely questioned. Perlman *et al.* (2014) conducted a study to find out if there is a positive impact of the working experience on the hazard recognition skills of construction superintendents. Although the aim is quite different than what it is questioned here, their results have remarkably common with these findings. They conducted experiments with civil engineering students, who have no working experience, construction superintendents and safety managers (Perlman *et al.*, 2014). They evaluated the number of correct hazards which identified by students and construction superintendents in the virtual environment (Perlman *et al.*, 2014). They adopted two-way between subject ANOVA test and found that students and construction superintendents performed similarly on hazard identification tasks (Perlman *et al.*, 2014). However, safety managers and students did not perform similarly on hazard recognition skills (Perlman *et al.*, 2014). They proposed two possible reasons behind these findings: (i) safety managers had taken more safety trainings compared with the construction superintendents and students or (ii) since their main role is to maintain the construction safety, therefore they performed better results (Perlman *et al.*, 2014). Similar to the study conducted by Perlman *et al*. (2014), the reason behind the difference in the performances between the students and engineers in this thesis is that the engineer subject group consists of engineers who have had many safety trainings and have an extensive knowledge on health and safety implementations. Since the difference stemmed from the fact that engineers have had already good performance on paper-based test before receiving V-SAFE v.2 compared with students' performances. However, V-SAFE v.2 still has significantly positive impact on the three groups' knowledge on the basic scaffolding installation tasks and related health and safety implementations. In the study conducted by Perlman et al. (2014), although the construction superintendents had received safety instructions beforehand, their performance have been associated with their safety culture. In their experiment, almost all the safety managers (85.7%) and construction superintendents

(92.8%) had received safety instruction while only 34.7% of the students had received safety instructions before (Perlman *et al*., 2014). Therefore, comparing safety performances of construction workers and engineers, safety culture might influence in their overall results.

In accounting research area, Trottier and Gordon (2018) focused on the appropriateness of using students as surrogates for managers and made an extensive literature review on this issue. They found that there is no robust conclusion about the appropriateness of students for managers (Trottier and Gordon, 2018). They suggested that the characteristics of students (e.g. work experience, gender, grade or knowledge) should be considered (Trottier and Gordon, 2018). For example, they found significantly differences between the responses of students and managers in accounting study, however, there are similarities between the response of students having work experience and the managers (Trottier and Gordon, 2018). Since the results on this issue provide mixed conclusions, there is no clear conclusion that students are always or never convenient surrogates for managers (Ashton and Kramer, 1980).

According to the results, the appropriateness of using students as surrogates depends on the research area, research scope and the demographic information of the participants in the experiments. In this case, the experiments have been conducted in Turkey. Meaning that, the results might change depending on the country and its own safety training implementations. Considering the research report by Eurofound and ILO in 2019, 45% of workers in the US, 41% in Europe have been reported that they have had training paid or provided by their companies over the last 12 months, while only 16% of workers have been reported in Turkey. Moreover, on-the-job training have been offered to 59% of workers in the US, 34% of workers in Europe, and 15% of workers in Turkey (Eurofound and ILO, 2019). Comparing these percentages, it can be said that if these experiments are conducted in the US or Europe, for construction workers, the initial safety performances before receiving the virtual safety training might be better comparing with the safety performances tracked in this thesis.

In this thesis, students are found as suitable surrogates for construction workers in the development of virtual safety training tools. However, the students might not be appropriate surrogates for engineers.

5.3. Supplementary Findings

V-SAFE v.2 performances, its relationship between the demographic information of the participants and feedbacks from the participants regarding V-SAFE v.2 are investigated via graphical representations.

V-SAFE v.2 scores are generated at the end of the testing module. After completing the training module, participants are asked to take the testing module of V-SAFE v.2. In testing module, participants are expected to remember safe acts given in the training module. 13 safe acts are determined in the testing module of V-SAFE v.2. Therefore, if a participant completes the testing module without any mistake, the score is generated as 13. However, if they perform unsafe act in the tool, they experience the consequence of their unsafe act and their score is presented accordingly. For example, if a participant forgets to lock the pin in the crossing of the scaffold, their representative player falls, the training simulation ends, and V-SAFE v.2 result is generated on the computer screen. According to V-SAFE v.2 scores, the mean scores are found as: 11.50 for students, 12.50 for construction workers, 12.50 for engineers. In other words, students performed 88.46% success while construction workers and engineers performed 96.15% success on V-SAFE v.2 testing module.

The relationship between demographic information and V-SAFE v.2 performances have been investigated only for construction workers since only this subject group have diverse demographic information. On the other hand, the factors of demographic information have not been found as determinant factors among engineer and student subject groups. Demographic information of the participants is gathered via questionnaire form. The form covers the demographic information such as gender, age, computer gaming experience and the number of working years in their profession. The aim is to see if the safety performance is related one of those variables. For example, computer gaming experience might have a positive impact on the safety performances on V-SAFE v.2. Since all workers are male, the gender is not a meaningful variable to be tested. However, age, work experience and computer gaming experience can be evaluated. Demographic information with the safety performances are investigated through graphical representations.

In Figure 5.3, V-SAFE v.2 score for each construction worker is presented with the age of correspondent construction worker. Ages of construction worker are sorted increasingly so that the relationship between two variables can be seen clearly. The dashed line represents the trend line of safety performances of 20 construction workers in Figure 5.3. While age is sorted in an increasing manner from left to right, the safety performances do not follow the similar path. The trend line of safety performances is almost flat. Therefore, considering 20 construction workers, their safety performances do not depend on their ages.

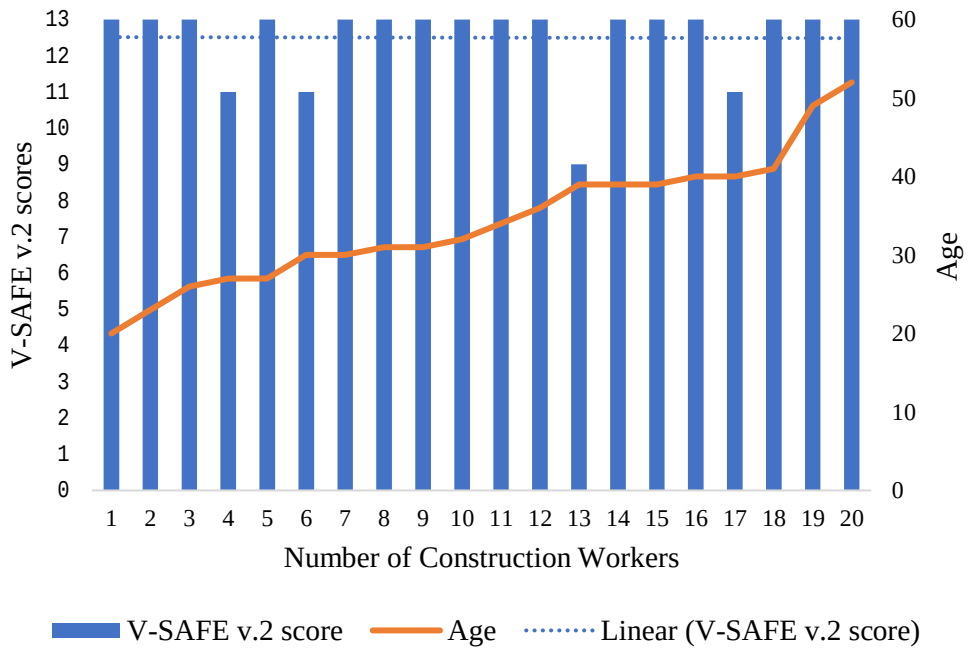

Figure 5.3. Age and V-SAFE v.2 Scores of Construction Workers.

Computer gaming experience of construction workers are also asked via questionnaire form. Construction workers are asked if they have ever played computer games before. If the answer is no, then they are categorized in “no computer gaming experience”. If the answer is yes, the next question investigates if they have ever played 3D computer games before. To make the question clearer, the most common 3D computer games are listed in the question. If the answer is no, but they also answered yes for the first question, then they are categorized in “2D computer gaming experience”. The last group called “3D computer gaming experience” is the ones who answered yes to both questions. The aim is to observe if the computer gaming experience of construction workers influences

their overall performance on V-SAFE v.2. Figure 5.4 represents the three groups of construction workers based on their computer gaming experience and their results on V-SAFE v.2. Results showed that construction workers with no computer gaming experience performed on V-SAFE v.2 without making any mistake. Meaning that, the computer gaming experience does not have an impact on construction workers' safety performance on V-SAFE v.2.

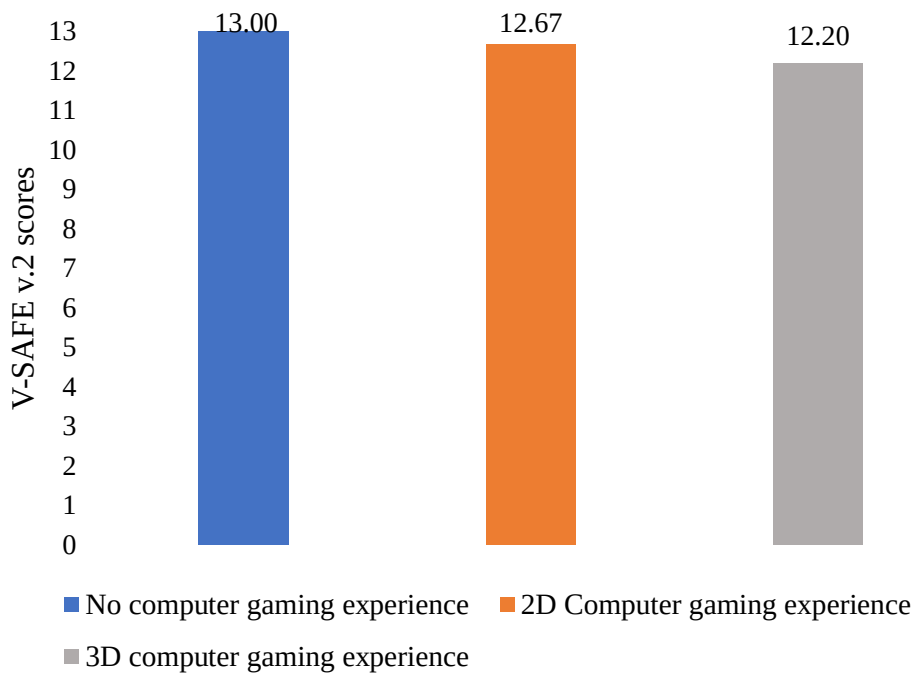

Figure 5.4. Mean V-SAFE v.2 scores of Construction Workers.

Additionally, the number of years spent in the construction industry has been asked to construction workers in a questionnaire form. Figure 5.5 illustrates the correlation between the work experience and V-SAFE v.2 scores of 20 construction workers. To make the illustration more understandable, the work experience is ranked as an increasing manner and corresponding V-SAFE v.2 scores presented as clustered columns. According to Figure 5.5, trend line (dashed line) for V-SAFE v.2 scores of construction workers and the orange line representing the work experience do not follow consistent paths. In other words, work experience does not have an impact on construction workers' safety performance on V-SAFE v.2.

Figure 5.5. Work Experience and V-SAFE v.2 Scores of Construction Workers.

Evaluations of the participants regarding V-SAFE v.2 are assessed via survey. A survey consists of 13 questions based on the Kirkpatrick model. This model includes four stages: reaction, learning, behavior and results (Kirkpatrick, 1998). The questions cover the three stages of the Kirkpatrick model: reaction for 1 – 4 questions, learning for 5 – 9 and behavior for 10 – 13 and results are feedbacks obtained from the participants. These questions are presented as 5-point Likert scale from 1 which stands for strongly disagree to 5 which stands for strongly agree. Participants were asked to evaluate different aspects of their learning experience in V-SAFE v.2 by completing a survey. The mean ratings of each question are shown in Figure 5.6 for three subject groups. Overall mean ratings for all questions are found above 4 which corresponds to mostly agree. Grouping the questions by the Kirkpatrick model, the overall mean ratings is 4.62 for reaction, 4.48 for learning and 4.66 for behavior. Overall ratings revealed that the participants agree that V-SAFE v.2 is considered as effective tool on participants' learning regarding the subjects covered in the tool. Some studies developing similar virtual safety training tools get positive feedback from the participants through surveys as well.

Figure 5.6. Mean Ratings given by Three Subject Groups.

The highest mean rating among the questions is that "Do you think that such trainings can be adopted for other educational activities?". In other words, participants agree that such high engaging trainings can be adopted in future studies regarding the educational activities. Le *et al.* (2015) conducted a survey based on 5-point Likert scale regarding the developed virtual reality system. They got positive feedback from the students and professionals that the system offers effective learning on construction safety (Le *et al.*, 2015).

5.4. Limitations and Future Directions

This thesis includes two limitations that might be addressed in future research. The first limitation is related to visuals in the virtual environment-based safety training tool due to economic considerations. The visuals do not fully represent the dynamic and complex nature of the construction environment. However, since the scope of the evaluation in the participants' learning covers only the basic scaffolding installation and related health and safety issues rather than all risks at the construction sites, the tool enables us to assess their relative performance on this scope.

Additionally, it is hard to find volunteering construction company and appropriate location to conduct the experiments. Therefore, the two construction companies involved in the experiments and provided volunteering construction workers and engineers for the experiments. The number of construction workers and engineers is adequate to conduct the experiments and test the hypotheses. However, it has been assumed that the subject groups of construction workers and engineers from two construction companies represent their all colleagues in the construction industry.

V-SAFE v.2 might be improved with the help of funds for future research so that experiments can be conducted with better quality visuals. Moreover, the subject groups of construction workers and engineers might be selected from more diverse construction projects in future research. In this way, it would be more representative for all construction workers and engineers. V-SAFE v.2 covers the basic scaffolding tasks and corresponding health and safety implementations. Further studies may extend the scope of the tool by adding more tasks in a broader sense.

6. CONCLUSION

Construction safety is a globally major concern by considering high rates of occupational accidents and fatalities worldwide. Although improvements in safety regulations have been made recently, safety problems are still high in the construction industry. Root causes of occupational accidents and preventive measures have been investigated in many research areas. Among the determinant factors to construction safety, safety trainings are found as one of the most powerful factors for the construction workers. Considering this, researchers has begun to question the effectiveness of current safety trainings on construction workers' safety knowledge and skills. Most construction safety trainings consist of classroom lectures and safety booklets which do not totally engage construction workers and the quality of traditional safety trainings is found as ineffective which leads to lack of safety knowledge among the construction workers. In this purpose, virtual environments have been recently integrated in the development of safety training tools for construction workers. Recent studies showed that safety trainings developed in virtual environment are more effective on workers' learnings compared with traditional classroom trainings.

In the development stages of virtual safety trainings tools, although the target groups are mostly construction workers or engineers rather than students, students are mainly used as subject group to test the tool due to time, location and money considerations. The use of students as surrogates for construction workers and engineers in the development safety training tools is commonly adopted in construction research areas, however, its appropriateness has not been fully investigated. Therefore, the aim of this thesis is to investigate the appropriateness of using students as surrogates for construction workers and engineers in the development of virtual safety training tools. Newly developed V-SAFE v.2 training tool was adopted in the experiments to investigate the aim and objectives of this thesis. This tool covers the basic scaffolding tasks coupled with related health and safety implementations at the construction site. The effectiveness of V-SAFE v.2 has also been evaluated in students, construction workers and engineer subject groups separately.

Experiments have been conducted with 20 civil engineering students, 20 construction workers and 20 engineers. Experimental settings have been designed to conduct the experiments in an effective manner. Experiment materials consist of: mini questionnaire, 10-question paper-based test covering basic information of scaffolding installation tasks and relative health and safety implementations given in V-SAFE v.2, the virtual environment-based safety training tool called V-SAFE v.2 training and testing modules and, (iv) a survey based on the Kirkpatrick model. Mini questionnaire has been administered to gain insight about the participants' demographic information such as age, work experience and computer gaming experience. Paper-based tests have been administered before and after V-SAFE v.2. Paper-based tests given prior to V-SAFE v.2 represents initial knowledge of the participants. In this way, the change in their knowledge regarding the topics covered in V-SAFE v.2 can be investigated quantitatively. In the last stage of the experimental setting, participants have been asked to complete a 13-question survey in order to collect their feedback regarding V-SAFE v.2. Analysis methods used in this thesis are listed as: Paired T-test to test the effectiveness of V-SAFE v.2 on students, Wilcoxon signed-rank test to test the effectiveness of V-SAFE v.2 on construction workers and engineers and two-factor mixed ANOVA to investigate if students can be used as surrogates for construction workers and/or engineers. Statistical analyses have been carried out in SPSS statistical package.

The effectiveness of V-SAFE v.2 on three subject groups have been investigated by comparing pre-test and post-test scores of the participants. Results revealed that V-SAFE v.2 training module has significantly positive effect on learning regarding the basic scaffolding knowledge and related health and safety implementations for three subject groups. These results are in line with similar studies on virtual safety training tools on various kinds of subjects in the literature.

Regarding the appropriateness of using students as surrogates for construction workers, results revealed that students and construction workers showed similar safety performances on the paper-based tests. When it comes to the appropriateness of using students as surrogates for engineers, safety performances have been found significantly different between two subject groups. In other words, students and engineers did not perform similar results on the paper-based tests. Consequently, combining two results, students can

be used as surrogates for construction workers. However, students might not be appropriate surrogates for engineers in the development of virtual safety training tools.

Moreover, participants performances on V-SAFE v.2 have been generated at the end of testing a result page. There are 13 steps that the participants are expected to accomplish. According to V-SAFE v.2 scores, students performed 88.46% success while construction workers and engineers performed 96.15% success on V-SAFE v.2 testing module. Overall results show that all subject groups have accomplished V-SAFE v.2 with high performances.

Correlation between safety performances of construction workers on V-SAFE v.2 and demographic information has been evaluated as supplementary findings. This correlation has been investigated only for construction workers because only construction workers have diverse demographic information in terms of age, work experience and computer gaming experience. According to the graphical representations, safety performances of construction workers do not depend on the age, working experience and computer gaming experience. This reveals that, the developed V-SAFE v.2 can be easily used by construction workers with diverse backgrounds.

Feedbacks from the participants regarding V-SAFE v.2 have been assessed by administering 13-question survey based on the Kirkpatrick model. Overall mean ratings for all questions have been found as above 4 which corresponds to mostly agree in five-point Likert scale. The highest mean rating among the questions is that “Do you think that such trainings can be adopted for other educational activities?”. Meaning that, participants have agreed on that these virtual environment-based trainings can be adopted in future studies regarding the educational activities. Overall ratings revealed that the participants agree that V-SAFE v.2 is considered as effective tool on participants’ learning regarding the subjects covered in the tool.

This thesis aims to enlighten the appropriateness of use of students as surrogates for construction workers and engineers in the development of virtual safety training tools. To fulfil this aim, the V-SAFE training tool has been adopted and used in the experiments. In the scope of this thesis, it has revealed that students can be used as surrogates for construction workers in the development of such safety training tools. However, students might not be

good surrogates for engineers in this sense. Findings are presented based on the subject groups in this thesis. Therefore, the subject size can be extended and selected from diverse construction projects in order to validate the results in a broader sense. Additionally, the virtual safety training tool might be further developed in order to enhance the visuals and extend the scope by adding more tasks. Therefore, further studies would be necessary to evaluate how these findings can be extended and generalized in an effective manner.

REFERENCES

Abdelhamid, T. S. and J. G. Everett, 2000, "Identifying root causes of construction accidents.", *Journal of Construction Engineering and Management*, Vol. 126, No. 1, pp. 52-60.

Abdel-Khalik, A. R., 1974, "On the efficiency of subject surrogation in accounting research." *The Accounting Review*, Vol. 49, No. 4, pp. 743-750.

Albert, A., M. R. Hallowell, and B. M. Kleiner, 2014, "Experimental field testing of a real-time construction hazard identification and transmission technique.", *Construction Management and Economics*, Vol. 32, No. 10, pp. 1000-1016.

Ashton, R. H. and S. S. Kramer, 1980, "Students as surrogates in behavioral accounting research: Some evidence." *Journal of Accounting Research*, pp. 1-15.

Awwad, R., O. El Souki and M. Jabbour, 2016, "Construction safety practices and challenges in a Middle Eastern developing country.", *Safety science*, Vol. 83, pp. 1-11.

Becker, P. and J. Morawetz, 2004, "Impacts of health and safety education: Comparison of worker activities before and after training." *American journal of industrial medicine*, Vol. 46, No. 1, pp. 63-70.

Bhoir, S. and B. Esmaeili, 2015, "State-of-the-art review of virtual reality environment applications in construction safety." In *Architectural Engineering National Conference, 2015*, pp. 457-468.

Bureau of Labor Statistics, 2017, "Current Population Survey, Census of Fatal Occupational Injuries", US Department of Labor.

Bureau of Labor Statistics, 2018, "Distribution of nonfatal occupational injuries and illnesses by private industry sector, 2017" US Department of Labor, https://www.bls.gov/iif/soii-chart-data-2017.htm, November 2018.

Burke, M. J., S. A. Sarpy, K. Smith-Crowe, S. Chan-Serafin, R. O. Salvador and G. Islam, 2006, "Relative effectiveness of worker safety and health training methods." *American journal of public health*, Vol. 96, No. 2, pp. 315-324.

Carter, G. and S. D. Smith, 2006, "Safety hazard identification on construction projects." *Journal of construction engineering and management*, Vol. 132, No. 2, pp. 197-205.

Chen, A., M. Golparvar-Fard and B. Kleiner, 2013, "Design and development of SAVES: A construction safety training augmented virtuality environment for hazard recognition and severity identification." In *Computing in Civil Engineering*, pp. 841-848.

Choudhry, R. M., D. Fang and S. Mohamed, 2007, "The nature of safety culture: A survey of the state-of-the-art." *Safety science*, Vol. 45, No. 10, pp. 993-1012.

Choudhry, R. M. and D. Fang, 2008, "Why operatives engage in unsafe work behavior: Investigating factors on construction sites." *Safety science*, Vol. 46, No. 4, pp. 566-584.

Çomu S., 2018, "Sanal Ortam Tabanlı İşçi Sağlığı Ve İş Güvenliği Simülasyon Aracının Geliştirilmesi", Project No. 315M186, The Scientific and Technological Research Council of Turkey (3501).

Dale, E., 1969, *Audio-visual methods in teaching*, 3rd Edition, A Holt-Dryden Book, Henry Holt and Company, New York.

Dickinson, J. K., P. Woodard, R. Canas, S. Ahamed and D. Lockston, 2011, "Game-based trench safety education: development and lessons learned." *Journal of Information Technology in Construction*, Vol. 16, No. 8, pp. 119-134.

Dytham, C., 2011, *Choosing and using statistics: a biologist's guide*, 3rd Edition, John Wiley & Sons, Chichester.

Edward, L., D. Lourdeaux, J. P. A. Barthès, D. Lenne and J. M. Burkhardt, 2008, "Modelling autonomous virtual agent behaviours in a virtual environment for risk." *International Journal of Virtual Reality*, Vol. 7, No. 3, pp. 13-22.

Elliott, A. C. and W. A. Woodward, 2007, *Statistical analysis quick reference guidebook: With SPSS examples*, Sage, London.

Elliott, W. B., F. D. Hodge, J. J. Kennedy and M. Pronk, 2007, "Are MBA students a good proxy for nonprofessional investors?" *The Accounting Review*, Vol. 82, No. 1, pp. 139-168.

Eurofound and International Labour Organization, 2019. "Working conditions in a global perspective". *Publications Office of the European Union, Luxembourg, and International Labour Organization*, Geneva.

Eurostat, 2018, "Accidents at work statistics", https://ec.europa.eu/eurostat/statistics-explained/pdfscache/11539.pdf , January 2020.

Gao, Y., V.A. Gonzalez, and T. W. Yiu, 2019, The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review. *Computers & Education*, Vol. 138, pp. 101-115.

Ghasemi, A., and S. Zahediasl, 2012, "Normality tests for statistical analysis: a guide for non-statisticians." *International journal of endocrinology and metabolism*, Vol. 10, No. 2, p. 486.

Guo, H., H. Li, G. Chan and M. Skitmore, 2012, "Using game technologies to improve the safety of construction plant operations." *Accident Analysis & Prevention*, Vol. 48, pp. 204-213.

Gürcanlı, G. E. and U. Müngen, 2013, “Analysis of construction accidents in Turkey and responsible parties.” *Industrial health*, Vol. 51, No.6, pp. 581-595.

Haslam, R. A., S.A. Hide, A. G. Gibb, D. E. Gyi, T. Pavitt, S. Atkinson and A. R. Duff, 2005, “Contributing factors in construction accidents.” *Applied ergonomics*, Vol. 36, No.4, pp. 401-415.

Health and Safety Executive, 2019, “Construction Statistics in Great Britain 2019”, UK http://www.hse.gov.uk/statistics/industry/construction.pdf

Health and Safety Executive, 2019. Workplace fatal injuries in Great Britain 2018, UK http://www.hse.gov.uk/statistics/pdf/fatalinjuries.pdf

Helander, M. G., 1991, “Safety hazards and motivation for safe work in the construction industry.” *International Journal of Industrial Ergonomics*, Vol. 8, No. 3, pp. 205-223.

Hu, K., H. Rahmandad, T. Smith-Jackson and W. Winchester, 2011, “Factors influencing the risk of falls in the construction industry: a review of the evidence.” *Construction Management and Economics*, Vol. 29, No. 4, pp. 397-416.

Hughes, C. T. and M.L. Gibson, 1991, “Students as surrogates for managers in a decision-making environment: An experimental study.” *Journal of Management Information Systems*, Vol. 8, No. 2, pp. 153-166.

Irizarry, J. and D. M. Abraham, 2005, “Application of virtual reality technology for the improvement of safety in the steel erection process.” In *Computing in Civil Engineering* pp. 1-11.

Karaca, N. G. and B. Gökçek, 2015, “Employers’ Occupational Health and Safety Training Obligations in Framework Directive and Training Procedure and Rules in Turkey.” *International Journal of Medical and Health Sciences*, Vol. 9, No. 3, pp. 860-863.

Kassem, M., L. Benomran and J. Teizer, 2017, “Virtual environments for safety learning in construction and engineering: seeking evidence and identifying gaps for future research.” *Visualization in Engineering*, Vol. 5, No.1, pp. 16.

Kerr, A. W., H. K. Hall and S. A. Kozub, 2002, *Doing statistics with SPSS*, Sage, London.

Killingsworth, J., T. Stentz, A. Moser and J. Grundman, 2014, “Effective instructional methods for providing safety training to construction workers.” In *Associate Schools of Construction: Proceedings of the 50th ASC Annual International Conference*.

Kirkpatrick, D. L., 1998, “The four levels of evaluation.” In *Evaluating corporate training: Models and issues*, Springer, Dordrecht, pp. 95-112,

Le, Q. T., A. Pedro and C. S. Park, 2015, “A social virtual reality-based construction safety education system for experiential learning.” *Journal of Intelligent & Robotic Systems*, Vol. 79, No. 3-4, pp. 487-506.

Li, H., G. Chan and M. Skitmore, 2012, “Visualizing safety assessment by integrating the use of game technology.” *Automation in construction*, Vol. 22, pp. 498-505.

Li, H., M. Lu, G. Chan and M. Skitmore, 2015, “Proactive training system for safe and efficient precast installation.” *Automation in Construction*, Vol. 49, pp. 163-174.

Lin, K. Y., J. W. Son and E. M. Rojas, 2011, “A pilot study of a 3D game environment for construction safety education.” *Journal of Information Technology in Construction (ITcon)*, Vol. 16, No. 5, pp. 69-84.

Lokka, I. E. and A. Çöltekin, 2020, Perspective switch and spatial knowledge acquisition: Effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. *Cartography and Geographic Information Science*, Vol. 47, No. 1, pp. 14-27.

Messner, J. I. and M. Horman, 2003, “Using advanced visualization tools to improve construction education.” In *Proceedings of Construction Applications of Virtual Reality (CONVR) 2003 Conference*, pp. 145-155.

Milgram, P. and F. Kishino, 1994, A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, Vol. 77, No. 12, pp. 1321-1329.

Mohd, N. I., K. N. Ali, S. Bandi and F. Ismail, 2019, “Exploring gamification approach in hazard identification training for Malaysian construction industry.” *International Journal of Built Environment and Sustainability*, Vol. 6, No. 1, pp. 51-57.

Mortensen, T., R. Fisher and G. Wines, 2012, “Students as surrogates for practicing accountants: Further evidence.” In *Accounting Forum*, Taylor & Francis, Vol. 36, No. 4, pp. 251-265.

Murrar, S. and M. Brauer, 2018, “Mixed model analysis of variance.” *The SAGE encyclopedia of educational research, measurement, and evaluation*, pp. 1075-1078.

Namian, M., A. Albert, C. M. Zuluaga and M. Behm, 2016, “Role of safety training: Impact on hazard recognition and safety risk perception.” *Journal of construction engineering and management*, Vol. 142, No. 12, p. 04016073.

Pan, Z., A. D. Cheok, H. Yang, J. Zhu, and J. Shi, 2006, “Virtual reality and mixed reality for virtual learning environments.” *Computers & graphics*, Vol. 30, No. 1, pp. 20-28.

Pedro, A., Q. T. Le and C. S. Park, 2016, “Framework for integrating safety into construction methods education through interactive virtual reality.” *Journal of Professional Issues in Engineering Education and Practice*, Vol. 142, No. 2, p. 04015011.

Perlman, A., R. Sacks and R. Barak, 2014, “Hazard recognition and risk perception in construction.” *Safety science*, Vol. 64, pp. 22-31.

Reid, K. and D. Dawson, 2001, “Comparing performance on a simulated 12-hour shift rotation in young and older subjects.” *Occupational and Environmental Medicine*, Vol. 58, No. 1, pp. 58-62.

Robson, L., C. Stephenson, P. Schulte, B. Amick, S. Chan, A. Bielecky and R. Peters, 2010, “A systematic review of the effectiveness of training & education for the protection of workers.” *Toronto: Institute for Work & Health*, p. 127.

Sacks, R., A. Perlman and R. Barak, 2013, “Construction safety training using immersive virtual reality.” *Construction Management and Economics*, Vol. 31, No.9, pp. 1005-1017.

Sawacha, E., S. Naoum and D. Fong, 1999, “Factors affecting safety performance on construction sites.”, *International journal of project management*, Vol. 17, No.5, pp. 309-315.

Shi, Y., J. Du, E. Ragan, K. Choi and S. Ma, 2018, “Social influence on construction safety behaviors: a multi-user virtual reality experiment.” In *Construction Research Congress*, pp. 174-183.

Shin, M., H. S. Lee, M. Park, M. Moon and S. Han, 2014, “A system dynamics approach for modeling construction workers’ safety attitudes and behaviors.” *Accident Analysis & Prevention*, Vol. 68, pp. 95-105.

Shuptrine, F. K., 1975, “On the validity of using students as subjects in consumer behavior investigations.” *The Journal of Business*, Vol. 48, No. 3, pp. 383-390.

Social Security Institution, 2017, “2017 Statistical Yearbooks” *SGK Yayını*.

Tixier, A. J. P., M. R. Hallowell, A. Albert, L. Van Boven and B. M. Kleiner, 2014, “Psychological antecedents of risk-taking behavior in construction.” *Journal of Construction Engineering and Management*, Vol. 140, No.11, p. 04014052.

Trottier, K. and I. M. Gordon, 2018, “Students as surrogates for managers: Evidence from a replicated experiment.” *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, Vol. 35, No.1, pp. 146-161.

Van Buren, M. E. and W. Erskine, 2002, “State of the industry: ASTD’s annual review of trends in employer-provided training in the United States.” *Alexandria, VA: American Society for Training and Development*.

Walpole, R. E., R. H. Myers, S. L. Myers and K. Ye, 2012, *Probability and Statistics for Engineers and Scientists*, 9th Edition, Prentice Hall, New York.

Wilkins, J. R., 2011, “Construction workers’ perceptions of health and safety training programmes.” *Construction Management and Economics*, Vol. 29, No. 10, pp. 1017-1026.

Xie, H., E. Tudoreanu and W. Shi, 2006, “Development of a virtual reality safety-training system for construction workers.” *Digital library of construction informatics and information technology in civil engineering and construction*.

Zhao, D. and Y. Ye, 2012, “Using virtual environments simulation to improve construction safety: an application of 3D online-game based training.” In *Future control and automation*, pp. 269-277.

APPENDIX A: SURVEY FORM

1) **Strongly disagree**
2) **Disagree**
3) **Neither agree nor disagree**
4) **Agree**
5) **Strongly agree**

		1	**2**	**3**	**4**	**5**
1	Do you agree that the dangers shown in the virtual training tool represent real situations at the construction sites?					
2	Do you think that the time you have spent during the virtual safety training was efficient?					
3	Do you agree that the virtual safety training is an entertaining experience?					
4	Do you think that this kind of trainings can be adopted for other training activities?					
5	Do you think the information you have learned in the virtual safety training is understandable and permanent?					
6	Do you agree that the virtual safety training has increased your knowledge on the scaffolding installation and related work safety issues?					
7	Do you think you were concentrated during the virtual safety training tool?					
8	Do you think that the virtual safety training tool experience was consistent with your experience in the scaffolding installation?					
9	Do you agree that the training you have received is a qualified training?					
10	Do you think that you will remember the information you learnt in the virtual safety training tool in the future?					
11	Do you recommend similar trainings to your friends?					
12	Do you want to receive a similar training in the future?					
13	Do you think that the training you received will help prevent accidents at the construction site?					

APPENDIX B: ASSUMPTION CHECKS FOR PAIRED T-TEST

Table B.1. Test of Normality for Paper-based Test Scores of Students.

Shapiro-Wilk			
Variables	Statistic	df	Sig.
Pre_Test_Score	0.931	20	0.162
Post_Test_Score	0.929	20	0.147

Table B.2. Test of Normality for Paper-based Test Scores of Construction Workers.

Shapiro-Wilk			
Variable	Statistic	df	Sig.
Pre_Test_Score	0.918	20	0.0923
Post_Test_Score	0.904	20	0.0496

Table B.3. Tests of Normality for Paper-based Test Scores of Engineers.

Shapiro-Wilk			
Variable	Statistic	df	Sig.
Pre_Test_Score	0.882	20	0.019
Post_Test_Score	0.846	20	0.005

APPENDIX C: ASSUMPTION CHECKS FOR TWO-FACTOR MIXED ANOVA

Table C.1. Test of Normality of Paper-based Test Scores in Hypothesis 4.

Shapiro-Wilk			
Variables	Statistic	df	Sig.
Pre_Test_Score	0.931	40	0.018
Post_Test_Score	0.915	40	0.005

Table C.2. Skewness and Kurtosis Values of Paper-based Test Scores in Hypothesis 4.

Variables		Statistic	Std. Error
Pre_Test_Scores	Skewness	0.221	0.374
	Kurtosis	-0.408	0.733
Post_Test_Scores	Skewness	0.017	0.374
	Kurtosis	-0.950	0.733

Table C.3. Levene's Test of Equality in Hypothesis 4.

Variable		Levene Statistic	df1	df2	Sig.
Pre_Test_Score	Based on Mean	1.520	1	38	0.225
	Based on Median	1.274	1	38	0.266
	Based on Median and with adjusted df	1.274	1	38	0.266
	Based on trimmed mean	1.534	1	38	0.223
Post_Test_Score	Based on Mean	1.764	1	38	0.192
	Based on Median	1.621	1	38	0.211
	Based on Median and with adjusted df	1.621	1	38	0.211
	Based on trimmed mean	1.776	1	38	0.191

Table C.4. Test of Normality of Paper-based Test Scores in Hypothesis 5.

Shapiro-Wilk			
Variable	Statistic	df	Sig.
Pre_Test_Score	0.948	40	0.062
Post_Test_Score	0.895	40	0.001

Table C.5. Skewness and Kurtosis Values of Paper-based Test Scores in Hypothesis 5.

Variable		Statistic	Std. Error
Pre_Test_Score	Skewness	-0.210	0.374
	Kurtosis	-0.178	0.733
Post_Test_Score	Skewness	-0.540	0.374
	Kurtosis	-0.118	0.733

Table C.6. Levene's Test of Equality in Hypothesis 5.

Variable		Levene Statistic	df1	df2	Sig.
Pre_Test_Score	Based on Mean	0.811	1	38	0.373
	Based on Median	0.798	1	38	0.377
	Based on Median and with adjusted df	0.798	1	34	0.378
	Based on trimmed mean	0.872	1	38	0.356
Post_Test_Score	Based on Mean	2.396	1	38	0.13
	Based on Median	2.652	1	38	0.112
	Based on Median and with adjusted df	2.652	1	37	0.112
	Based on trimmed mean	2.694	1	38	0.109