A MODEL TO EVALUATE THE EFFECTIVENESS OF VIRTUAL CONSTRUCTION SAFETY TRAINING TOOL

by

Işık Ateş Kıral

B.S., Civil Engineering, Istanbul University, 2011

M.S., Construction Project Management, The University of Manchester, 2012

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Civil Engineering

Boğaziçi University

2019

A MODEL TO EVALUATE THE EFFECTIVENESS OF VIRTUAL CONSTRUCTION SAFETY TRAINING TOOL

APPROVED BY:

Assist. Prof. Semra Çomu Yapıcı .....
(Thesis Advisor)

Prof. Hüsnü Murat Günaydın .....

Assoc. Prof.Beliz Özorhon Orakçal .....

Assoc. Prof. Esin Ergen Pehlevan .....

Assist. Prof. İnci Ayhan .....

DATE OF APPROVAL: 21.10.2019

ACKNOWLEDGEMENTS

I want to thank my invaluable thesis advisor, Assist. Prof. Semra Çomu for supporting and guiding me over the years. I also want to express my appreciation for her invaluable and constructive suggestions during the planning and development of the research. Her precious support, patience, willingness, and expertise in construction safety management research made the completion of study possible. She has been a great model of an excellent researcher, mentor, and instructor.

I would like to thank my committee members Prof. Hüsnü Murat Günaydın, Assoc. Prof. Beliz Özorhon Orakçal, Assoc. Prof. Esin Ergen Pehlevan, Assist. Prof. İnci Ayhan, for their valuable time, constructive comments, and recommendations.

I want to thank all firm members of Bertuğ Bey, Turner, Zetaş; who took part in the conducted experiments. Their positive and supportive attitudes and willingness to take part in the experiments are quite valuable to complete the experiments.

My special thanks also go to my wife, mom, dad, for their unconditional love. They have always supported and motivated me to overcome difficult times.

I also would like to thank all my family members, my mother in law, and my brother in law, who are the supporters of me all the time.

I want to extend my gratitude to my friends in Boğaziçi University; Semih Çağlayan, Sevilay Demirkesen Çakır, Ahmet Talha Gezgin, Onur Kalan, Gürkan Günay, Uğur Dündar, Emrah Kılıç, Hasan Altun, who never refrain from supporting and encouraged to finish this research. I would also like to present my special thanks to my friends; Emre Karakurt, Kadir Ağma, Kahraman Kasımoğlu, Kutay Bolkol, Selçuk Sönmez, Serkan Evyapan, who always supported me with their friendship, and motivated me in difficult times.

I should also be noted that this thesis is produced based on the research project funded by Boğaziçi University Research Fund (BAP) coded 17A04D5 and 14A0SUP1.

ABSTRACT

A MODEL TO EVALUATE THE EFFECTIVENESS OF VIRTUAL CONSTRUCTION SAFETY TRAINING TOOL

Construction projects have become more complex. Due to this complexity, construction projects are exposed to more risks and the probability of occupational accidents on the site increases. So, the provision of safety on the construction field becomes more challenging, and the number of occupational accidents and fatalities remains high all around the world. The literature highlight that human-related factors play a crucial role in the safety management process, so, improving behavior-based skills of workers by adequate safety training method has become important to provide on-site safety. Previous literature states that an effective safety training method should cover (i) information transfer, (ii) instructional feedback, (iii) knowledge development, (iv) hands-on practice, (v) behavioral modeling and (vi) hazard identification. Due to the low level of engagement, conventional safety training methods fail to provide an effective safety training. In order to bridge this gap, a high level of engagement between the safety training method and the participant should be provided. In this regard, the main aim of this study is to quantitatively analyze the effectiveness of virtual construction safety training tool entitled Virtual Safety Analysis For Engineering applications (V-SAFE). During evaluation, eye tracking technologies have been also used to test the attention level and safety awareness of participants. Analysis results show that V-SAFE provides all the essential features of effective safety training. The present study makes a significant contribution to the field of safety training, since it is the first and only study that evaluates all safety training parameters rather than focusing on a single parameter. It has also developed a conceptual model that examines the effectiveness of virtual safety training methods. This model offers great insights for future studies on how to measure the effectiveness of safety training.

ÖZET

SANAL İNŞAAT İŞ GÜVENLİĞİ EĞİTİM ARACININ ETKİSİNİN DEĞERLENDİRİLMESİ İÇİN BİR MODEL

İnşaat projeleri daha karmaşık hale gelmektedir. Bu karmaşık yapı sebebiyle, inşaat projeleri daha fazla risk içermekte ve şantiyelerdeki iş kazaları olasılığı artmaktadır. Bu nedenle, inşaat sahalarında etkin bir iş güvenliğinin sağlanması zorlaşmakta ve tüm dünyada inşaat sektöründe iş kazalarının ve ölümlerin sayısı artmaktadır. Literatür, insan odaklı faktörlerin iş güvenliği sürecinde çok önemli bir rol oynadığını göstermektedir, bu nedenle, çalışanların davranış temelli becerilerini etkin bir iş güvenliği eğitimi ile geliştirmek, sahada güvenliği sağlamak için önemli bir unsur olmuştur. Literatür, etkili bir iş güvenliği eğitimi yönteminin (i) bilgi aktarımı, (ii) eğitim geri bildirimi, (iii) bilgi birikimi gelişimi, (iv) aktif uygulama, (v) davranış modellenmesi ve (vi) tehlike tanımlaması unsurlarını kapsaması gerektiğini belirtmiştir. Düşük etkileşim seviyesinden ötürü geleneksel yöntemlerle etkin bir iş güvenliği eğitimi sağlanamamaktadır. Bu eksiğin giderilmesi için, iş güvenlik eğitim yöntemi ile katılımcı arasında daha yüksek bir etkileşim sağlanmalıdır. Bu bilgilerin ışığında, bu çalışmanın amacı, Virtual Safety Analysis For Engineering applications (V-SAFE) adlı iş güvenliği eğitim aracının etkinliğini, nicel olarak analiz etmektir. Ölçüm sırasında, katılımcıların dikkat seviyesini ve güvenlik farkındalığını ölçmek için göz izleme teknolojileri de kullanılmıştır. Analiz sonuçları, V-SAFE'in etkin güvenlik eğitiminin tüm temel özelliklerini sağladığını göstermiştir. Tek bir unsura odaklanmak yerine, tüm etkin iş güvenliği eğitimi unsurlarını değerlendiren ilk ve tek çalışma olduğundan, mevcut çalışma iş güvenliği eğitimi alanında literatüre önemli bir katkı sağlamaktadır. Ayrıca, sanal iş güvenlik eğitim yöntemlerinin etkinliğini ölçümünü inceleyen kavramsal bir model geliştirilmiştir. Bu model, sanal iş güvenliği eğitimi yöntemlerinin etkinliğinin nasıl ölçüleceği adına sonraki çalışmalar için faydalı görüşler sunmaktadır.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT v
ÖZET vi
LIST OF FIGURES xi
LIST OF TABLES xv
LIST OF SYMBOLS xix
LIST OF ACRONYMS/ABBREVIATIONS xx
1. INTRODUCTION 1
1.1. Problem Determination 2
1.2. Background of the Research 3
1.3. Aims and Objectives of the Research 4
1.4. Research Method 6
1.5. Research Limitations 7
1.6. Research Scope 8
2. LITERATURE REVIEW 10
2.1. Accident Types 10
2.2. Root Causes of the Accidents 11
2.3. The Importance and Effectiveness of Safety Training 13
2.4. Virtual Construction Safety Training 15
2.5. Review of the Mixed Reality Tools in the Construction Safety Training 18
2.5.1. Virtual Environment 30
2.5.2. Virtual Reality and Augmented Virtuality 33
2.5.3. Augmented Reality 34
2.6. Measuring the Effectiveness of Safety Training 36
2.6.1. Eye Tracking 38
3. RESEARCH METHODOLOGY 41
3.1. Hypotheses and Statements of the Research 41
3.1.1. Statement and Hypotheses Related to Information Transfer . . 41

3.1.2. Statement and Hypotheses Related to Instructional Feedback, Knowledge Development 44
3.1.3. Statement and Hypotheses Related to Hazard Identification . . 47
3.1.4. Statement and Hypotheses Related to Hands-On Practice . . . 49
3.1.5. Additional Related Statement and Hypotheses 51
3.2. Tools and Parameters to Use 53
3.2.1. Virtual Environment: V-SAFE (Virtual Safety Analysis For Engineering applications) 53
3.2.1.1. Input Devices and Gaming Controllers 58
3.2.2. Eye Tracking Devices 61
3.3. Data Analysis Methods 64
3.3.1. Comparing One Group for Two Related Samples 64
3.3.1.1. Paired Samples T-test 64
3.3.1.2. Wilcoxon Signed-Rank Test 65
3.3.1.3. McNemar Test 65
3.3.2. Comparing One Group for Several Related Samples 66
3.3.2.1. One-Way ANOVA with Repeated Measures 66
3.3.2.2. Friedman Test 67
3.3.3. Comparing Two Groups for Independent Samples 67
3.3.3.1. Student T-test 67
3.3.3.2. Mann-Whitney U test 68
3.3.4. Factorial Analysis 68
3.3.4.1. Two-Way ANOVA 68
3.3.4.2. Kruskal Wallis-H Test 69
3.3.4.3. Scheirer-Ray-Hare Test 69
3.3.5. Tests to Evaluate Assumptions 70
3.3.5.1. Shapiro-Wilk and Kolmogorov-Smirnov Test for Normality 70
3.3.5.2. Levene and Bartlett Tests for Homogeneity 71
3.3.5.3. Wald-Wolfowitz Runs Test for Randomness 71
3.3.5.4. Mauchly Test for Sphericity 72
3.3.5.5. Multivariate Test of Normality 72

3.3.5.6. Cohen t-test for the Power and Effect Size Evaluation 73
3.4. Experimental Design 75
3.4.1. Information Comprehensibility Experiment 76
3.4.2. Repeatable Virtual Experiment 79
3.4.3. Hazard Identification Experiment 81
3.4.4. Eye-Movement Behavior Comparison Experiment 84
3.4.5. Construction Site and Gaming Experience Experiment 86
4. ANALYSIS RESULTS 89
4.1. Information Comprehensibility Experiment 89
4.1.1. Descriptive Analyses 93
4.1.1.1. Testing the Power 93
4.1.1.2. Testing the Randomness 93
4.1.2. Analysis Results 94
4.2. Repeatable Virtual Experiment 95
4.2.1. Descriptive Analyses 96
4.2.1.1. Testing the Sphecirity of the Data 96
4.2.1.2. Testing the Effect Size, Power and Multivariate Normality 97
4.2.1.3. Testing the Randomness 98
4.2.2. Analysis Results 99
4.3. Hazard Identification Experiment 100
4.3.1. Descriptive Analyses 105
4.3.1.1. Testing the Power 105
4.3.1.2. Testing the Randomness 107
4.3.2. Analysis Results 109
4.4. Eye-Movement Behavior Comparison Experiment 110
4.4.1. Descriptive Analyses 111
4.4.1.1. Testing the Normality of the Data 112
4.4.1.2. Testing the Homogeneity of the Data 112
4.4.1.3. Testing the Randomness 113
4.4.1.4. Testing the Effect Size and Power 115
4.4.2. Analysis Results 115
4.5. Construction Site and Gaming Experience Experiment 117

4.5.1. Descriptive Analyses 119
4.5.1.1. Testing the Normality of the Data 119
4.5.1.2. Testing the Equality of Error Variances 120
4.5.2. Analysis Results 122
4.5.2.1. Kruskal-Wallis H Test with Scheier-Ray-Hare Extension 122
4.5.2.2. Testing the Effect Size 123
5. DISCUSSION 124
5.1. Introduction 124
5.2. Research Motivation 124
5.3. Safety Information Transfer 127
5.4. Instructional Feedback Mechanism and Knowledge Development 132
5.5. Hazard Identification 135
5.6. Hands-On Practice and Eye Tracking 139
5.7. Gaming and Site Experience 143
6. CONCLUSIONS 146
6.1. Key Findings of the Research 146
6.2. Revisiting the Research Aim, Objectives and Major Contribution of Thesis149
6.3. Limitations and Recommendations for Further Research 151
6.4. ACKNOWLEDGMENT 154
REFERENCES 155
APPENDIX A: THE EYE MOVEMENT BEHAVIOR EXPERIMENT RESULTS 179
APPENDIX B: THE QUESTION SHEET OF CRANE OPERATOR ROLE . 182
APPENDIX C: THE QUESTION SHEET OF POINTER ROLE 184
APPENDIX D: THE QUESTION SHEET OF BRICKLAYER ROLE 186

LIST OF FIGURES

Figure 1.1. Research Scope and Organization. 9

Figure 2.1. Reality-Virtuality Continuum (Milgram and Kishino, 1994). 19

Figure 2.2. The systematic review of the mixed reality technologies 1. 27

Figure 2.3. The systematic review of the mixed reality technologies 2. 27

Figure 2.4. The systematic review of the mixed reality technologies 3. 28

Figure 2.5. The systematic review of the mixed reality technologies 4. 28

Figure 2.6. The systematic review of the mixed reality technologies 5. 28

Figure 2.7. The systematic review of the research orientation 1. 29

Figure 2.8. The systematic review of the research orientation 2. 29

Figure 2.9. The systematic review of the research orientation 3. 30

Figure 3.1. Conceptual Framework of Data and Quality (Wang and Strong, 1996). 42

Figure 3.2. Hypothesis Related to Information Transfer. 44

Figure 3.3. Hypotheses Related to Knowledge Development and Instructional Feedback. 47

Figure 3.4. Hypotheses Related to Hazard Identification. 48

Figure 3.5. Hypotheses Related to Hands-On Practice. 51

Figure 3.6. Additional Hypotheses Related to Effective Safety Training. 52

Figure 3.7. An Avatar with PPE. 54

Figure 3.8. PPE Selection. 54

Figure 3.9. Visual risk sources integrated into V-SAFE. 55

Figure 3.10. V-SAFE Training Framework. 56

Figure 3.11. System Architecture of V-SAFE. 57

Figure 3.12. The interface of the crane operator in a real tower crane. 59

Figure 3.13. The controller setting of trainees during V-SAFE training. 59

Figure 3.14. Logitech F310. 60

Figure 3.15. The interaction of the trainees with V-SAFE via Logitech F310. . 60

Figure 3.16. Tobii X2-30. 63

Figure 3.17. Tobii Pro Glasses 2. 63

Figure 3.18. The flow chart of data analysis method selection for paired samples experiments. 74

Figure 3.19. The flow chart of data analysis method selection for independent samples experiments. 75

Figure 3.20. Gathering eye-tracking metrics through Tobii X2-30 and Tobii Pro Glasses 2. 84

Figure 3.21. The sample assignment of the area of interests on real construction site and V - SAFE. 85

Figure 4.1. The Appropriate Data Analysis Method for Information Comprehensibility Experiment. 92

Figure 4.2. The Appropriate Data Analysis Method for Repeatable Virtual Experiment. 99

Figure 4.3. The Appropriate Data Analysis Method for Hazard Identification Experiment for Paired Differences. 103

Figure 4.4. The Appropriate Data Analysis Method for Hazard Identification Experiment for Independent Groups. 104

Figure 4.5. The Appropriate Data Analysis Method for Eye-Movement Behavior Comparison Experiment. 114

Figure 4.6. The Appropriate Data Analysis Method for Construction Site and Gaming Experiment. 121

Figure 5.1. Conceptual Model to Evaluate the Effectiveness of Virtual Construction Safety Training Methods. 126

Figure 5.2. The illustration of four-hinge link system in Power Point presentation. 129

Figure 5.3. The illustration of weight load and dimensions in Power Point presentation. 132

Figure 5.4. The trigger effect of risk identification through visualizing warnings. 137

Figure A.1. The Eye Movement Behavior Experiment Results 1. 179

Figure A.2. The Eye Movement Behavior Experiment Results 2. 180

Figure A.3. The Eye Movement Behavior Experiment Results 3. 181

Figure B.1. The Question Sheet of Crane Operator Role 1. 182

Figure B.2. The Question Sheet of Crane Operator Role 2. 183

Figure C.1. The Question Sheet of Pointer Role 1. 184

Figure C.2. The Question Sheet of Pointer Role 2. 185

Figure D.1. The Question Sheet of Bricklayer Role 1. 186

Figure D.2. The Question Sheet of Bricklayer Role 2. 187

LIST OF TABLES

Table 2.1. Distribution of the accident types. 11

Table 2.2. Summary of the "Engagement Hypothesis" (Burke *et al.*, 2006). . . 14

Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training. 21

Table 3.1. The summary of eye movement metrics. 62

Table 3.2. Levenes' test values and corresponding effect size (adopted from Cohen, 2013). 74

Table 3.3. Experimental setting of the information comprehensibility experiment. 78

Table 3.4. Experimental setting of the repeatable virtual experiment. 80

Table 3.5. Experimental setting of the hazard identification experiment. . . . 83

Table 3.6. Experimental setting of the eye-movement behavior comparison experiment. 86

Table 3.7. Experimental setting of the construction site and gaming experience experiment. 88

Table 4.1. Information Comprehensibility Experiment Results. 89

Table 4.2. Information Comprehensibility Experiment Results based on Information Quality Parameters 90

Table 4.3. Categorization of Questions. 91

Table 4.4. Experiment and Data Features for Paired Samples. 92

Table 4.5. The Summary of the Power Test. 93

Table 4.6. The summary of randomness test analysis results 93

Table 4.7. Wilcoxon Signed-Rank Test Results. 94

Table 4.8. The status of the hypotheses. 94

Table 4.9. Analysis results of the information comprehensibility experiment. . 95

Table 4.10. Experiment and Data Features. 96

Table 4.11. Mauchly Test of Sphericity Results. 97

Table 4.12. Multivariate Test Results. 97

Table 4.13. The summary of randomness test analysis results. 98

Table 4.14. Post-Hoc Analysis Results. 100

Table 4.15. The status of hypotheses. 100

Table 4.16. Experiment Results of Hazard Identification Experiment. 101

Table 4.17. Experiment and Data Features for Paired Samples. 102

Table 4.18. Experiment and Data Features for Mutually Exclusive Groups. . . 104

Table 4.19. The Summary of the Power Test for 1st Group. 105

Table 4.20. The Summary of the Power Test for 2nd Group 105

Table 4.21. The Summary of the Power Test for the Improvement Between Two Independent Groups. 106

Table 4.22. Randomness Test Results of Group 1. 107

Table 4.23. Randomness Test Results of Group 2. 108

Table 4.24. The summary of McNemar test analysis results for 2nd Group. . . 109

Table 4.25. The status of the hypotheses. 110

Table 4.26. Experiment Results of Eye-Movement Behavior Comparison Experiment. 110

Table 4.27. Experiment and Data Features. 111

Table 4.28. Kolmogorov Smirnov Test Results. 112

Table 4.29. The summary of data variables. 112

Table 4.30. The variables descriptive and the results of Levenes' test of Homogeneity. 113

Table 4.31. The summary of randomness test analysis results. 113

Table 4.32. The summary of Cohen t-test analysis results 115

Table 4.33. Mann Whitney-U test analysis results. 116

Table 4.34. The status of the hypotheses in eye-movement behavior comparison. 116

Table 4.35. Analysis results of the trainees based on construction site experience.117

Table 4.36. Descriptive results. 118

Table 4.37. Experiment and Data Features. 119

Table 4.38. Test of normality results. 120

Table 4.39. Levene Test of equality of error variances. 120

Table 4.40. Kruskal-Wallis Ranks. 122

Table 4.41. Kruskal-Wallis Test Results. 122

Table 4.42. Achieved Effect Size. 123

Table 4.43. The status of the hypotheses in the construction site and gaming experience experiment. 123

LIST OF SYMBOLS

β	Probability of Type II Error

LIST OF ACRONYMS/ABBREVIATIONS

3D	Three Dimensional
ANOVA	Analysis of Variance
AOI	Area of Interest
ARCTM	Accident Root Causes Tracing Model
BIM	Building Information Modeling
BLS	Bureau of Labour Statistics
EEG	Electroencephalogram
EU	European Union
Eurostat	European Statistical Office
GDP	Gross Domestic Product
GPS	Global Positioning System
HSE	Health and Safety Executive
MRI	Magnetic Resonance Imaging
NPC	Non-Playable Character
OSHA	Occupational Safety and Health Administration
PC	Personal Computer
PPE	Personal Protective Equipment
RFID	Radio Frequency Identification
SSI	Social Security Institution
TSI	Turkish Statistical Institute
UK	United Kingdom
UNSD	United Nations Statistics Division
US	United States
UWB	Ultra-Wideband
V-SAFE	Virtual Safety Analysis for Engineering Applications

1. INTRODUCTION

The report of World Economic Forum states that the construction sector constitutes around 6% of world gross domestic product (GDP) in 2018 (World Economic Forum, 2018), while it is forecasted to reach 14.7% by 2030 (Oxford Economics, 2015). In addition, Renz and colleagues (2016) reported that, 1% improvement in the productivity around the world could achieve savings around $100 billion in one year term, when it potentially contributes to the competitiveness of country and sustainable development (Schwab, 2017; Despotovic *et al.*, 2016).

The construction sector has an essential role in the economy of any country. In the European Union (EU), the construction industry accounts for around 9% of gross domestic product (GDP) and offers 18 million employment opportunities, which makes construction a strategically important sector for the European countries (European Commision, 2016). Similarly, in the United States (US), the construction industry is worth around $781.4 billion, which corresponds to more than 4% of the US national economy in 2017 (US Bureau of Economic Analysis, 2018). Moreover, construction employment is around 10 million, which corresponds to 16.7% of the US workforce in 2016 (US Census Bureau, 2017). In Turkey, the construction sector is worth around 517 billion Turkish Liras, while, the current GDP of Turkey is about 6298 billion Turkish Liras in 2017, that corresponds to 8.2% of Turkish GDP (Turkish Statistical Institution (TSI), 2018a). In 2017, there were around 2 million employees existed in the Turkish construction sector, which corresponds to 7% of employment in Turkey (TSI, 2018b). In brief, the previous reports highlight that the construction sector is one of the leading industries that changed the global world economy. Employment opportunities and large-budget construction projects are essential facts showing the impact of the construction sector on the world economy. Therefore, the world economy could be affected by many factors that influence the construction sector (Özorhon, 2012).

Due to the significant impacts of the construction industry, such as improving GDP, increasing employment, the demand for the construction industry has been continuously growing. Growing demand for the construction industry usually requires the adaptation of high technologies (Briesemeister, 2018). Therefore, construction projects get more innovative and complex (Briesemeister, 2018) and effective management of construction projects has become increasingly challenging (Pinto and Morris, 2014). As a result of the complexity, the uncertainty level of the construction projects is relatively high compared to other industries (Hobday, 1998).

1.1. Problem Determination

Construction projects have continuously become more complex and innovative (Briesemeister, 2018). Due to this complexity, construction projects are exposed to more risks and the probability of occupational accidents increases (Briesemeister, 2018). For this reason, the provision of safety on the construction field becomes more challenging, and the number of occupational accidents and fatalities remains high in the construction industry all around the world. For instance, in Turkey, 1633 fatal accidents and 587 deaths were recorded in the construction industry in 2017 (Social Security Institution (SSI), 2019a), which corresponds to 17.5% of all recorded occupational accidents and 36% of fatalities in Turkey (SSI, 2019a). According to the SSI statistics, the construction sector ranks first with the mining and metal sectors in terms of the number of occupational accidents in Turkey (SSI, 2019b). Besides, it is seen that the industry constitutes one-third of the total deaths, and the sector is in the first place with these ratios (SSI, 2019b). Similarly, in 2015, a total of 3876 fatal occupational accidents occurred in 28 EU countries (European Statistical Office (Eurostat, 2018a) and approximately 20% of these fatal accidents occurred in the construction sector (Eurostat, 2018b). Catastrophically, the construction industry has the highest fatal accident rate in the 28 EU countries (Eurostat, 2018b). Just as EU countries and Turkey, In the US, 991 fatal accidents were recorded in the construction industry in 2016 (Bureau of Labour Statistics (BLS), 2017). Compared to other sectors, the construction industry had the highest number of occupational deaths in 2016 (BLS, 2017). In brief, construction field employees have been subject to severe accidents at work

and achieving a safe working environment is an ongoing concern in the construction industry all around the world, including developed countries.

Based on the high number of fatal accidents in the construction industry, construction companies require unique and novel methods to decrease the number of on-site occupational accidents. The previous literature about construction safety (e.g. Abdelhamid and Everett, 2000; Huang and Hinze, 2003; Shapira and Lyachin, 2009) concentrates on safety performance on construction sites in terms of reducing the number of accidents and mitigating the impact of accidents. For instance, Abdelhamid and Everett (2000) identified the root causes of accidents on construction sites. They (Abdelhamid and Everett, 2000) found that one of the primary reasons for accidents is the unsafe behavior of the workers. Huang and Hinze (2003) analyzed the falling from height accidents, which represents the highest rate leads to fatal injuries in the United States. Huang and Hinze (2003) found that the improper hazard identification of the workers is the most significant cause of accidents. Shapira and Lyachin (2009) identified the leading factors to provide safety on construction sites during tower crane activities. Shapira and Lyachin (2009) found that the safety attitudes of the crane operators play a crucial role in preventing on-site accidents. As evidenced by these examples, the literature indicates the significance of human-based safety factors in preventing on-site accidents.

1.2. Background of the Research

To improve the human-based safety skills, safety training is considered as one of the critical strategies (e.g., Cohen *et al.*, 1998; Demirkesen and Arditi, 2015; Wong *et al.*, 1999; Tam and Fung, 2011). On the other hand, conventional construction safety training methods fail to improve the behavior-based safety skills, since these methods (e.g., lectures, videos, textbooks, etc.) only focus on transferring the safety information (Burke *et al.*, 2006). However, effective safety training should establish a base for trainees to improve their safety behavior by appropriate hands-on practice (Burke *et al.*, 2006). In other words, the previous literature (Burke *et al.*, 2006; Albert *et al.*, 2014) state that an effective construction safety training method should contain

both: i) information transfer, ii) knowledge development, iii) hands-on training, iv) behavioral modeling, v) programmed instructional feedback, vi) hazard identification. Due to the complex nature of the construction projects, on-site safety training is not feasible, and it even increases the risk of an accident especially for the inexperienced new workforce (Chi *et al.*, 2005). For providing an effective construction safety training method, virtual construction safety training could be an appropriate solution.

To provide adequate safety training, recent literature (e.g., Guo *et al.*, 2013; Perlman *et al.*, 2014; Lin *et al.*, 2011) has focused on virtual construction safety training methods. For example, Guo and colleagues (2013) developed a novel framework for virtual prototyping technology to support the safety management process of construction projects. Analysis results demonstrate that virtual prototyping techniques considerably improved safety performance by reducing the safety management cost by 30%, without having any fatal accidents in Tseung Kwan O Sports Ground project in Hong Kong (Guo *et al.* 2013). Similarly, another recent study by Perlman and colleagues (2014) explored to what extent construction supervisors could comprehend the hazardous situations on the construction site and how they assess the risks. The analysis results show that superintendents who toured a virtual construction site using a 3-sided virtual reality model identified a higher number of risks, compared to subjects used conventional risk identification methods such as reviewing photographs and construction documents (Perlman *et al.*, 2014). Similarly, Lin and colleagues (2011) developed a virtual environment based working environment for the students to improve their hazard recognition skills. The results demonstrated that the virtual environment has a positive impact on the students' hazard recognition skills. Taken together, virtual methods have a high potential to ensure a safe workplace in the construction industry.

1.3. Aims and Objectives of the Research

As evidenced by the previous studies related to construction safety training, the literature highlights the importance of the virtual tools and techniques in providing adequate safety training in the construction industry. As mentioned, the previous literature (Burke *et al.*, 2006; Albert *et al.*, 2014) state that an effective construction

safety training method should cover both: i) information transfer, ii) knowledge development, iii) hands-on training, iv) behavioral modeling, v) programmed instructional feedback and vi) hazard identification. Therefore, high number of studies (Cheng and Teizer, 2013; Bosché *et al.*, 2015; Perlman *et al.*, 2012; Guo *et al.*, 2012; Lin *et al.*, 2011; Hadikusumo and Rowlinson, 2004; Le *et al.*, 2015; Park and Kim, 2013; Zhao and Lucas, 2015) focused on the development of an effective safety training method by integrating some parameters such as information transfer, hazard identification, feedback, etc. However, these studies focused on only one or two of specific parameters of the efficient safety training. For instance, the research conducted by Lin and colleagues (2011) put attention to improve hazard identification level of the workers. However, this study did not focus on providing hands-on training and knowledge development. Similar to Lin and colleagues' study, all studies did not focus on all essential parameters of practical safety training. The main challenge in providing an adequate safety training method is associated with the complexity of evaluating the effectiveness of the training method. Therefore, the majority of previous studies are primarily used the conventional performance measurement techniques, such as collecting user feedbacks via interviews, conducting paper-based examinations and closed-ended questions or checking the trainees' number of identified hazards. These methods could be able to evaluate the safety information level, hazard identification performance; however, they are not adequate to assess the safety knowledge and safety behavior of trainees. In order to bridge this gap, a virtual construction safety training method is developed, which automatically evaluates the safety performance of trainees via hands-on practice. Thus, it could be able to analyze safety performance in an unbiased quantitative manner. In addition, eye-tracking methods are involved during the evaluation of safety behavior. The previous literature (Henderson *et al.*, 2013) highlights that eye-tracking methods could be suitable to provide an objective analysis of the behavior patterns of the trainees. In brief, there are two significant gaps existed in the literature. Previous studies did not focus on all parameters of effective safety training. Also, these studies merely used conventional performance analysis methods, which fail to evaluate knowledge development and hands-on practice effectiveness. In order to bridge this gap, the main aim of this study is to quantitatively analyze the effectiveness of virtual construction safety training tool entitled Virtual Safety Analysis For Engineering applications

(V-SAFE) in terms of all parameters of effective safety training. Followings are the objectives of the research:

- Measuring the efficiency of the developed virtual construction safety training method: V-SAFE.
- Evaluating construction workers' behavior patterns through eye-tracking methods.
- Gathering relevant feedback about the use of V-SAFE from the workers; the main target group.
- Developing a conceptual model that explains the effectiveness of virtual construction safety training methods.

1.4. Research Method

A virtual safety training method (V-SAFE) is developed under BAP Grant No. 7902. The validity of V-SAFE in terms of safety training effectiveness is analyzed through several statistical analysis methods as Wilcoxon Signed Rank Test, One-Way ANOVA with Repeated Measures, McNemar Test, Mann-Whitney U Test, Kruskal Wallis-H Test with Scheirer-Ray-Hare extension. In addition, several descriptive analysis methods as Shapiro-Wilk Test, Kolmogorov-Smirnov Test, Levene Test of Equality, Wald-Wolfowitz Runs Test, Mauchly Test, Multivariate Test of Normality and Cohen Test are used to evaluate, whether the datasets provide the assumptions of the aforementioned statistical analysis methods. During the experiments, several data sets are gathered through; i) data analysis via automated performance analysis procedure of advanced virtual tool, ii) the paper-based examinations results of trainees and iii) the eye-tracking metrics gathered through Tobii eye-tracking devices. The developed virtual tool and two different eye-tracking devices (mobile, and screen-based) were used in the experiments.

1.5. Research Limitations

The scope of this research focuses on the measurement of safety training tool effectiveness that involves significant parameters of training. Behavioral modeling is an essential parameter for effective safety training. To evaluate the effectiveness of the safety training method based on behavioral modeling, it is required to assess the trainees' behavior over a long period. For instance, if it is aimed to evaluate the performance of safety training based on trainees' behavior, it is vital to monitor and assess the changes in trainees' safety behavior. However, due to the time limitation of the research and impossibility of entering a real construction site daily, it is not possible to directly evaluate the effectiveness of the behavior modeling level. To minimize this limitation, the adjustment of safety behavior is analyzed through feedback interventions and knowledge development, which are essentials of safety behavioral modeling (Burke *et al.*, 2006). Moreover, by using eye-tracking methods, the snapshot of safety behavior patterns of workers was obtained. Although the longitudinal data from workers could not be gathered, it was gained insight into the safety behavior of the workers, through eye-tracking metrics.

Another limitation of the research is the subjects used in some of the experiments. There are concerns about the use of students as subjects in construction health and safety training research. A limitation of conducting experiments using students as trainees had been recognized long before (Lorge *et al.* 1958). However, there is still a considerable amount of literature (e.g. Bosché *et al.*, 2015; Chen *et al.*, 2016; Chi *et al.*, 2015; Ding *et al.*, 2013; Hadipriono and Barsoum, 2002; Juang *et al.*, 2013; Lu and Davis, 2016; Paes *et al.*, 2017; Pedro *et al.*, 2015; Perlman *et al.*, 2014; Teizer *et al.*, 2013; Le *et al.*, 2015) utilize students as surrogates due to the difficulty in persuading construction companies to take apart in the experiments about construction safety process. Consequently, the studies related to construction safety training use, graduate and undergraduate students as the first step to validate the usability of the training procedures. To minimize the impact of not using industry professionals, students who completed the on-site internship have been used in this research. Hence, all trainees participated in the experiments were familiar with construction site activities. Also,

previous studies (e.g., Perlman *et al.*, 2014) show that there are not any significant differences between the performance of students and industry professionals. For instance, Perlman and colleagues (2014) compared the hazard identification level of students and superintendents. Analysis results show that there is not any statistically significant difference between the hazard recognition level of students and the superintendents. In light of these findings, in this study, graduate and undergraduate students were used as subjects in four of the five conducted experiments. However, as the most significant experiment that evaluates the level of hands-on training level, eye-movement behavior experiment is conducted through construction workers.

1.6. Research Scope

This research involves six chapters. The first stage is the introduction of this research, which summarizes this thesis by describing the main elements, such as research aim, objectives, scope, and limitations. In the second stage, a detailed, systematic, comprehensive literature review is conducted, to involve research motivation and represent substantive findings, theoretical and practical contributions of previous literature about construction safety training. Moreover, earlier studies about virtual construction safety training methods including eye-tracking research are indicated. In the third stage, the methodology of the research and the development of the virtual safety training tool are explained. In addition, relevant literature about statistical testing methods and the experimental procedure is represented. Also, developed hypotheses and statements are described in this section. In the fourth stage, analysis results of all experiments are presented. As the fifth stage, findings of the study are discussed in the context of efficient safety training. Moreover, the findings of the research are compared with the previous studies in the literature. In the sixth stage, conclusion and recommendations for future studies are described. Figure 1.1 represents the scope and organization of the research.

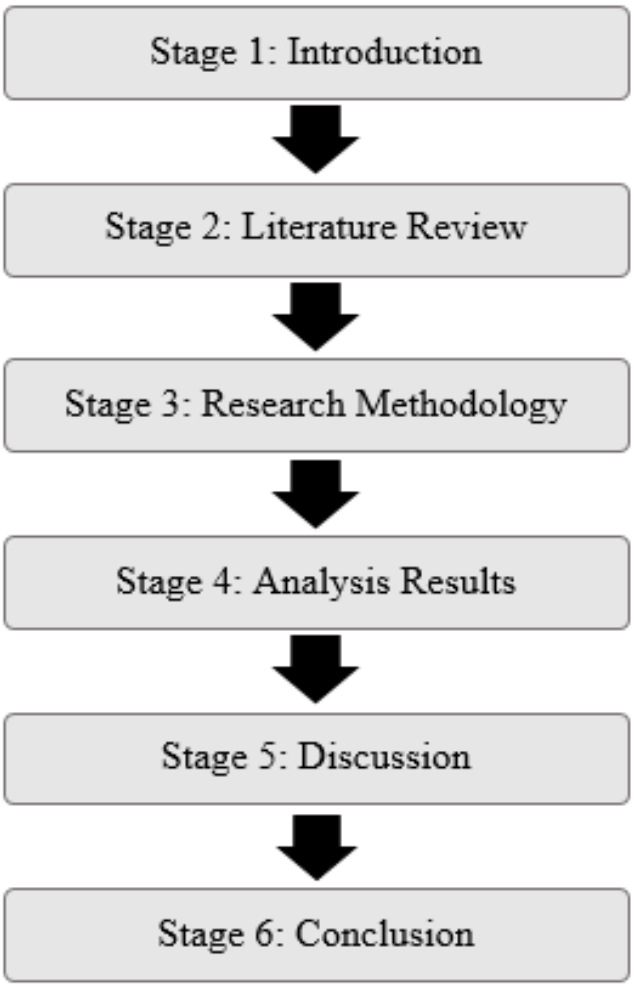


Figure 1.1. Research Scope and Organization.

2. LITERATURE REVIEW

Construction projects have a complex structure which leads to an increase in the probability of occupational accidents (Briesemeister, 2018). Consequently, providing a safe work environment is quite difficult. Accordingly, the number of occupational accidents and fatalities remains high in the construction industry all around the world. Compared to other sectors, the construction industry had the highest number of occupational deaths in many countries such as US, EU and Turkey 2016 (Eurostat, 2018b; SSI, 2019a; SSI, 2019b; BLS, 2017). In brief, construction field employees have been subject to severe accidents at work and achieving a safe working environment is an ongoing concern in the construction industry all around the world, including developed countries.

2.1. Accident Types

To reduce accident rates, identifying frequently encountered severe accidents are an essential step. SSI investigated 2578 fatal occupational accidents between 2013 and 2017 to determine the causes of these accidents (2019b). SSI reports indicate that seven accident types lead to fatalities on the construction in Turkey (2019b). Similarly in the US, kind of fatal accidents that occurred in the construction sector has also similarities with Turkey (Goh and Ubernarayama, 2017). In a recent study, Goh and Ubeynarayana (2017) processed around 1000 labeled records of accidents existed in Occupational Safety and Health Administration (OSHA) reports and categorized these accidents and found ten fatal accident types occurred in the US construction industry. In the UK, five different types of accidents caused death (Health and Safety Executive (HSE), 2018). Consequently, there are a total of twelve different types of accidents causing fatalities in the construction sector (Table 2.1).

Table 2.1. Distribution of the accident types.

Accident Type	**Study / Institution**	**Country and Year**	**Frequency**
1 - Falling From Height	Goh and Ubeynarayana, 2017 HSE, 2018 SSI, 2019b	Turkey, 2013 - 2017 US, 2016 UK, 2018	Turkey - 44.4% US - 23.6% UK - 48.0%
2 - Electrocution	Goh and Ubeynarayana, 2017 HSE, 2018 SSI, 2019b	Turkey, 2013 - 2017 US, 2016 UK, 2018	Turkey - 6.0% US - 10.8% UK - 6.0%
3 - Soil and Structure Collapses	Goh and Ubeynarayana, 2017 HSE, 2018 SSI, 2019b	Turkey, 2013 - 2017 UK, 2018 US, 2016	Turkey - 3.5% US - 21.2% UK - 12.0%
4 - Struck by Falling or Moving Object	Goh and Ubeynarayana, 2017 HSE, 2018	US, 2016, 2017 UK, 2018	US - 17.7% UK - 11.0%
5 - Heavy Equipment Accidents	SSI, 2019b	Turkey, 2013 - 2017	Turkey - 5%
6 - Fires and Explosions	Goh and Ubeynarayana, 2017	US, 2016	US - 4.7%
7 - Traffic Accidents	Goh and Ubeynarayana, 2017 HSE, 2018 SSI, 2019b	Turkey, 2013 - 2017 US, 2016, 2017 UK, 2018	Turkey - 17.8% US - 6.3% UK - 9.0%
8 - Materials-based Accidents	SSI, 2019b	Turkey, 2013 - 2017	Turkey - 4.3%
9 - Exposure to Chemical Substances	Goh and Ubeynarayana, 2017	US, 2016	US - 2.9%
10 - Caught in / Between Objects	Goh and Ubeynarayana, 2017	US, 2016	US - 2.9%
11 - Exposure to Extreme Temperatures	Goh and Ubeynarayana, 2017	US, 2016	US - 1.7%
12- Health Problems	SSI, 2019b	Turkey, 2013 - 2017	Turkey - 10.4%
13 - Other	Goh and Ubeynarayana, 2017 SSI, 2019b	Turkey, 2013 - 2017 US, 2016	Turkey - 8.8% US - 4.3%

2.2. Root Causes of the Accidents

Higher numbers of occupational accidents in the construction industry indicate the inadequacies in the safety management process. Consequently, many researchers (e.g. Abdelhamid and Everett, 2000; Carter and Smith, 2006; Dester and Blockley, 1995; Hale *et al.*, 2012; Huang and Hinze, 2003; Lingard and Rowlinson, 2005; Patrucco *et al.*, 2010; Shapira and Lyachin, 2009; Tam *et al.*, 2004) have focused on the causes of the accidents occurred on the construction site. In their study, Abdelhamid and Everett (2000) developed an accident root causes tracing model (ARCTM) to meet the

safety requirements of the construction industry. Analytical results show that causes of the accidents are; i) not adequately identifying an unsafe condition before the activity was initiated, ii) continuing to proceed with a task after a worker recognizes a potentially unsafe condition and (iii) acting unsafely despite the high-risk conditions of the work environment. Moreover, Abdelhamid and Everett (2000) listed the causes of the unsafe conditions as i) the lack of management actions, ii) the unsafe behavior of the field employee, iii) force majeure events and iv) improper working conditions. Consistently, Carter and Smith (2006) analyzed the hazard identification level of workers in three construction projects in the UK. Study results indicate that 33% of the risks could not be identified (Carter and Smith, 2006). Moreover, Tam and colleagues (2004) identified the major accident causes of the unsafe conditions as i) lack of safety awareness of the field employee, ii) inadequate safety supervision and iii) poor safety training. Also, Dester and Blockley (1995) listed the primary cause of site accidents in construction projects is the inadequate safety culture, hence, improper safety behavior. In their study, Hale and colleagues (2012) analyzed the causes of the 26 previous accidents occurred in the British construction industry. Analysis results show that the unsafe acts and preconditions to dangerous acts are some of the major causes of these 26 accidents. Similarly, Patrucco and colleagues (2010) analyzed the accidents occurred on the construction sites in Italy. Study results show that approximately 90% of the risk sources could not be identified precisely in the Italian construction industry. Moreover, Huang and Hinze (2003) analyzed the significant causes of fall accidents, which is the most common accidents on construction sites. Huang and Hinze (2003) found that misjudgment of the hazardous situation is the most crucial reason for fall accidents. Similarly, Shapira and Lyachin (2009) analyzed the tower crane operations safety. Shapira and Lyachin (2009) found that the safety attitudes of the operators play a significant role in the provision of safety in crane activities. In a similar study, Lingard and Rowlinson (2005) stated that around 80 to 90% of the construction accidents are primarily or non-directly associated with the unsafe behavior of the workers. Taken together, inadequate human-based safety skills such as low hazard identification, improper safety attitudes, inappropriate safety behaviors, etc. lead to accidents on the construction sites. In that case, improving human-based safety skills plays a vital role in the provision of safety on the construction site.

2.3. The Importance and Effectiveness of Safety Training

In order to improve the human-based safety skills, the previous literature (e.g., Wilkins, 2011; Wirth and Sigurdsson, 2008; Held *et al.*, 2001) highlights the importance of effective safety training. Wilkins (2011) evaluated the conceptions of the workers against safety training. Analytical results show that safety training improves the knowledge of the employees when their perceptions are positive about the safety training method (Wilkins, 2011). Similarly, Wirth and Sigurdsson (2008) highlighted that safety training could change the behaviors of the trainees, and this behavior adjustment can promote health and safety. Moreover, Held and colleagues (2001) analyzed the use of safety training on the workers' skin protection. Held and colleagues (2001) found that proper safety training and ergonomics training improved both the safety behavior and safety knowledge level. In the light of this research, the construction safety training has a great potential to enhance the human-based safety skills factors such as safety attitudes, safety behaviors, safety knowledge, etc. of the field employee.

In order to improve the human-based safety skills, many studies focused on the effectiveness of the construction safety training in the literature (e.g. Kirkpatrick, 1979; Alliger and Janak, 1989; Frese and Zapf, 1994; Burke *et al.*, 2006). As one of the first known studies in the training effectiveness field, Kirkpatrick (1979) developed a training effectiveness framework. According to Kirkpatrick, an effective training program should involve reaction, learning, behavior, and results (1979). Kirkpatrick (1979:120) defined reaction as "how well the trainees liked a particular training program. Moreover, Kirkpatrick (1979:120) states, "evaluating in terms of reaction is the same as measuring the feelings of the conferees". Besides, learning refers to the "what principles, facts, and techniques were understood and absorbed by the conferees (Kirkpatrick, 1979:126). Behavior was defined as how the participants adjust their behaviors based on the training (Kirkpatrick, 1979:132). Finally, Kirkpatrick (1979:137) represented results as the level of the achieved examination results of the participants. However, Kirkpatrick assumed that the elements in the framework (reaction, learning, behavior, and results) are independent of each other. To bridge this gap, Alliger and Janak (1989) analyzed the correlation among the elements in Kirkpatricks' model (1979) and

found that there is a positive correlation between both learning and behavior, behavior and results, and learning and results. On the other hand, Alliger and Janak (1989) did not consider the effect of a training method to the effectiveness of the training procedure. In this context, Frese and Zapf (1994) argue that effective training method that enables more interaction, such as lectures, interviews are more superior compared to passive techniques such as written notes. In other words, Frese and Zapf (1994) state that the transmission of the relevant information by the more proactive approach and practicing play a crucial part in the effectiveness of the training. Similarly, Burke and colleagues (2006) developed a novel framework to evaluate safety training performance. The method proposed by Burke and colleagues (2006) could be defined as "engagement hypothesis" that assessed the training effectiveness based on the level of the trainee engagement. Burke and colleagues categorized the safety training methods into three categories such as: (i) least engaging, (ii) moderately engaging and (iii) highly engaging (2006). Least engaging methods have been defined as the methods that merely focus on the information transfer such as lectures, videos, texts, etc. In other words, these methods were criticized by Burke and colleagues (2006) as being inadequate in terms of not evaluating the safety knowledge level of the participants. In moderately engaging methods, programmed instructional feedback mechanism was integrated, together with information transfer. Most common, moderately engaging methods are computer-based instruction systems. Finally, most engaging training methods consist of information transfer, knowledge development, hands-on training, behavioral modeling, and instructional feedback mechanism (Table 2.2).

Table 2.2. Summary of the "Engagement Hypothesis" (Burke *et al.*, 2006).

	Least Engaging Methods	**Moderately Engaging Methods**	**Highly Engaging Methods**
Information Transfer	x		x
Programmed Instructional Feedback		x	x
Hands-On Practice			x
Knowledge Development			x
Behavioral Modelling			x

Moreover, Albert and colleagues (2014) state that hazard identification is one of the essential factors of effective construction safety training. Consequently, an effective construction safety training procedure should provide: i) information transfer,

ii) knowledge development, iii) hands-on training, iv) behavioral modeling, v) programmed instructional feedback, vi) hazard identification. Construction companies have frequently used information transfer based safety training methods that fail to cover the major features of effective training such as knowledge development, hands-on experience, hazard identification and getting feedback (Burke *et al.*, 2006). Therefore, an alternative construction safety training method is required to improve training effectiveness. In this case, virtual technologies provide a significant opportunity to advance the level of safety training.

2.4. Virtual Construction Safety Training

Due to the complex nature of the construction projects, on-site safety training is not feasible (Sisson, 2001) and it even increases the risk of an accident especially for the inexperienced new workforce (Li *et al.*, 2012a). Besides, possible errors of the employees during training could lead to delays in the project duration; hence, cost overrun. Therefore, active training during a real construction project becomes costly, hazardous, and impractical. Construction companies have frequently used only information transfer based least engaging training methods that fail to cover the significant features of training such as knowledge development, hands-on practice, behavioral modeling, and instructional feedback (Burke *et al.*, 2006). Consequently, an alternative construction safety training method that contains information transfer, knowledge development, hands-on training, behavioral modeling, programmed instructional feedback, hazard identification is required to improve safety training effectiveness. In this case, the virtual technologies-based simulation provides a significant opportunity to advance the level of the training.

A recent and growing body of literature has investigated the use of virtual safety training tools in the construction industry. In a comparative analysis, Albert *et al.* (2014) aim to improve the hazard recognition skills of the workers by using the virtual reality-based safety training tool. As a result, a human-computer interactive augmented virtuality training platform was developed, and cognitive mnemonics based on energy sources such as mechanical, electrical, etc. were presented. The results

show that identifying hazards via the energy-based cognitive mnemonics significantly improves the hazard identification level with the support of the systematic categorization. Similarly, Cheng and Teizer (2013) developed a novel framework that gathers the positional data of the field workers, thus, provides a holistic approach for safety training. Study results show that both construction and safety information could be transferred via visualized information; hence, several safety skills of the workers such as situational awareness, safety attitudes are significantly improved. In their study, Bosché and colleagues (2015) integrated a mixed reality-based safety training for the construction trade workers via the development of a 6-degree-of-freedom head pose tracking system. Analytical preliminary results reveal that head pose tracking systems are quite helpful for the trade workers to experience realistic construction site conditions, especially when they have worked at height. Similarly, another recent study by Perlman and colleagues (2014) explored to what extent construction supervisors could comprehend the hazardous situations on the construction site and how they assess the risks. The study results show that superintendents identified a greater number of risks compared to subjects used conventional risk identification methods such as photographs and documents (Perlman *et al.*, 2014). A study by Guo *et al.* (2012) integrated serious gaming technologies into the construction safety management process to improve training performance. The study reveals that the safety performance of the plant operatives and tower crane workers were significantly improved. Similarly, Guo and colleagues (2013) developed a novel framework for virtual prototyping technology to support the safety management process of construction projects. Moreover, comparative analysis results demonstrated that virtual prototyping techniques considerably improved the safety performance by reducing the safety management cost by 30% and without having any fatal accidents. Similarly, Lin and colleagues (2011) developed a virtual environment based working environment for the students to improve their hazard recognition skills. The results demonstrated that virtual environment has a positive impact on the students' hazard recognition skills. Lin and colleagues' (2011) contribution is supported by another study (Li *et al.* 2012b) that evaluates the tower crane dismantlement activities via virtual reality-based safety training. Moreover, a recent study by Kıral and colleagues (2015) examined the use of the virtual environment based construction safety training method for a tower crane operation

that provides hands-on safety training. Similarly, Hadikusumo and Rowlinson (2004) integrated a virtual reality construction model with its design for building a safety database, to support safety engineers to recognize safety hazards in the construction projects. Hadikusumo and Rowlinson showed that virtual reality technologies have a high potential to transfer tacit knowledge via visualization, hence, supporting safety engineers to identify hazard sources. Quite similarly, Le and colleagues (2015) developed a safety training tool for safety education purposes. The role-playing platform developed by Le and colleagues (2015) utilized dialogic and experiential learning for the students. Thus, students improved their knowledge about safety via hands-on practice. Moreover, Park and Kim (2013) created a framework for the visualization of the novel safety management process by integrating building information modeling (BIM), location tracking, and augmented reality. Study results show that safety management visualization systems improve the hazard identification level and real-time communication between the construction management team and workers. Zhao and Lucas (2015) developed a virtual reality-based environment which serves as a safe training environment for the workers and allows them to rehearse their task. Results show that the interactive virtual reality-based environment improves the cognitive abilities and risk awareness level of the participants. (Zhao and Lucas, 2015). Similarly, Li and colleagues (2012a) developed a 4D based visualization system to improve the safety knowledge and safety skills of the field employee. The study results show that the proposed system is useful in terms of improving trainees' safety knowledge and safety skills. In a recent study, Li and colleagues (2015a) developed a Proactive Construction Management System to improve safety skills and safety awareness. Study results show that proactive data acquisition method proposed by Li and colleagues (2015a) significantly improved the safety skills of the trainees. Talmaki and Kamat (2012) developed a framework to enable visualization of the construction workplaces via 3D geographic information. They (Talmaki and Kamat, 2012) aimed to develop an active warning system when the field employees reached the safety threshold. Study results show that developing a visualization based active warning system significantly improves the situational awareness of the field employees. Similarly, Teizer and colleagues (2013) integrated the real-time location data of construction workers to a virtual environment to improve their situational awareness via safety training. Study results show that

the integration of positional data improves the situational awareness of workers, thus, improves their safety skills. In a similar study, Hsiao and colleagues (2005) tested the impact of surround-screen virtual systems on the safety behaviors of the workers during scaffolding training. Hsiao and colleagues found that using surround-screen virtual systems improved the simulation level of the activities; thus, the safety behaviors, especially for the inexperienced workers. Similarly, Tichon and Diver (2010) analyze the impact of virtual simulation in terms of safety behavior improvement during plant operations. Tichon and Diver (2010) found that plant operators' performance was significantly improved via augmented virtuality based training. Similarly, Patrucco (2010) developed an approach about the computer image generation system for simulating construction activities to improve the hazard identification level of the trainees. Patrucco and colleagues (2010) found that the computer generation system significantly improves the visualization quality, hence, improves the hazard identification level of the trainees. Consequently, all the studies mentioned above show that mixed reality-based technologies have a high potential to provide effective safety training for construction projects.

2.5. Review of the Mixed Reality Tools in the Construction Safety Training

During the review of the literature about the construction safety training field, it has been recognized that the studies used (i) virtual reality, (ii) augmented reality, (iii) augmented virtuality and (iv) virtual environment. When the literature about the virtual tools are reviewed, it can be seen that all the virtual methods are the components of the reality-virtuality continuum defined by Milgram and Kishino (1994). In their review study, Milgram and Kishino (1994) described visual displays as a subset of the mixed reality technology. Milgram and Kishino (1994) state that mixed reality technology involves the major visual domains that provide integration between the real and virtual worlds. Therefore, they developed a continuum that explains the links between the elements in the virtual and real-world (Figure 2.1).



Figure 2.1. Reality-Virtuality Continuum (Milgram and Kishino, 1994).

Milgram and Kishino (1994) have considered virtual reality as an essential component of mixed reality. Virtual reality is defined by Squelch (2001: 210) as "3-D computer-generated representations of real or imaginary worlds with which a user can have real-time interaction and experience some feeling of being present in those worlds". The primary objective of using virtual reality is to enable the interaction by the provision of visual-based information flow (Hoffman, 1996; Champion, 2010).

Virtual environments are physical-digital workspaces that users could take place, interact with other users (Schröeder, 2008). In that case, Schröeder (2008:2) states the main difference between the virtual reality and the virtual environments as "against virtual worlds is that the latter term has been applied to persistent online social spaces; that is, virtual environments that people experience as ongoing over time and that have large populations which they experience together with others as a world for social interaction".

Augmented reality is another feature that have been used in the construction safety training literature. The main aim of the augmented reality is to enhance the information level of the objects existed in the real world; thus the reality level by the supplementation of the computer-generated information comes into users' experience (Wang and Dunston, 2005). Similarly, Azuma (1997) lists the primary characteristics of the augmented reality as (i) combination of the real and virtual world and (ii) interactivity in real-time. In the construction safety training literature, augmented reality is used, to improve the cognition level of the construction field employees by embedding the critical information about the safety issues.

Augmented virtuality denotes the connection tools that merge the virtual environment within real-world objects (Milgram and Kishino, 1994). Augmented virtuality is based on the integration of the physical objects into the digital workspace, which provides an opportunity to enhance the interaction level in the virtual environment (Silva and Sutko, 2009). As stated by Milgram and Kishino (1994), all the components of the mixed reality aim to improve the interaction level between the real and virtual world.

Taken into account all the recent studies in the literature, it can be concluded that virtual methods have a high potential to lead to a positive impact on the construction safety training process. On the other hand, the aims, scopes, and methods of the research show variability across the studies. With these ends in mind, a systematic literature review is represented to discuss the mixed reality components, system methodology, tools, interaction system, aim, scope, methodology, tools, research main focus, research outcomes, data gathering, and analysis methods and experiment population of mixed reality-based construction safety training studies. Thus, it becomes possible to evaluate the key factors that improve the effectiveness of virtual safety training in the construction projects by identifying the different training outcomes, hypotheses, and the propositions of the studies based on the elements involved in the reality-virtuality continuum. The research contributions of the studies are quite different from each other (Table 2.3). To develop a deeper understanding of the impact of virtual tools and research orientation, the systematic review of studies used mixed reality technologies in construction safety training is illustrated in Figure 2.2. Moreover, employed research methods of the studies are represented in Figure 2.3.

Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training.

Paper	Contribution
Albert *et al.* (2014, p.9)	"The developed SAVES training module can dramatically improve the hazard recognition skills of construction workers and potentially reduce construction injuries".
Albert *et al.* (2017, p.6)	"SAVES, revealing that work crews from the two projects were able to only identify 70% of gravity hazards in the baseline phase. Fortunately, the introduction of the intervention resulted in a positive impact and workers were able to identify 89% of hazards in the post-intervention phase". "Similar to findings from testing SAVES, introducing the SMQM maturity model as an intervention also resulted in statistically significant improvements in hazard recognition across all hazard types"
Behzadan and Kamat (2013)	"Students can also move their AR Books around the room to form groups, virtually manage a project, discuss a certain scenario, and explore alternative solutions in a collaborative setting, while learning basic concepts such as equipment operations, jobsite safety, resource utilization, work sequencing, and site layout".
Benjaoran and Bhokha (2010, p.401)	"The prototype could identify working-at-height hazards according to the progress of the construction project. Different building components were identified with the hazards at their own placements"
Bosché *et al.* (2015, p.10)	"The system illustrates the VR immersive environment that enables trainees to experience height, without involving any actual work. Ultimately, it could even enable them to start to accustom themselves to such conditions".
Chen *et al.* (2015, p.8)	"The alert views were the changing background color of the ambient views and the superposing of the attention signal onto the focused views; it was used to raise participants' attention levels to impending dangers".
Cheng and Teizer (2013, p.15)	"The view provided by the 3D display improved situational awareness of users and allowed views from multiple camera locations".
Chi *et al.* (2012, p.650)	"In evaluating the efficiency and effectiveness of the designed UIs, the result of the usability test shows that erection tasks conducted using multiple views and with virtual information guidance are more efficient than traditional operations".
Chi *et al.* (2015, p.446)	"Laser scanning devices can be used for producing as-built 3D models which can in turn be compared with as-planned BIM models for deviations checking GPS and RFID can help the material dispatching control, while automatic accessibility analysis can be used for solving dispatching conflicts
Ding *et al.* (2013, p.37)	"The system put much more emphasis on the integration of safety monitoring and early warning in real time than previously possible with conventional methods. Therefore, more potential accidents could be identified and prevented before they occurred".
Dzeng *et al.* (2016, p.66)	"The search pattern analysis results could provide valuable information for safety trainers and educators. Both the experienced and novice workers exhibited a high number of fixations on attention points involving ladders, implying that they require more time to determine whether these situations are hazards".
Fang and Cho (2016, p.66)	"Operators' Safety Awareness was assessed by an indirect, objective method that quantifies Safety Awareness by five Key Performance Indicators with respect to efficiency and safety performance. The proposed framework was implemented on a real crane and validated in a series of actual lift tasks. Test results indicate that the proposed framework was able to improve the safety performance (i.e., load clearance and load sway control) in different lift tasks, especially in blind lift scenarios. The results also imply that using the system has adverse impacts on lift time efficiency because the operators did not feel confident enough to fully rely on the system".
Golovina *et al.* (2016, p.114)	"The novelty of the presented method is that it is a first step in providing safety personnel with realistic and construction site-specific leading safety indicator data rapidly and related to the frequency, location, and contributing causes of near miss events".

Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training (cont.).

Paper	Contribution
Goulding *et al.* (2012, p.115)	"Cognisant of this, the main aim of the OSP VR interactive training environment was to help different stakeholders involved in the supply chain appreciate the nature and complexity of OSP projects, by allowing them to ?try out' OSP working practices in a safe and controlled learning environment. Hence, the VR interactive training environment can be considered as essential, as it purposefully allows learners to appreciate new methods, processes and thinking required".
Guo *et al.* (2012, p.211)	"Game technology-based training method provides trainees with an easily operated multi-user virtual environment that allows them to try and study different methods of operating the plant. At the same time, it provides trainees with sufficient knowledge of operations to help them identify potential safety problems".
Guo *et al.* (2013, p.22)	"The visual plant operation materials were offered to the operators. They could easily understand these potential safety problems, and practice their skills in operating construction plants in the virtual environment. Thus the safety problems were avoided during real construction".
Hadikusumo & Rowlinson (2004, p.288)	"Virtual reality can represent virtually real construction processes which also have safety hazards; and the tool is equipped with a possible safety hazard database and accident precaution database which can stimulate the engineers to recall their experience".
Hadipriono and Barsoum (2002, p.136)	"With this tool, trainees are expected to appreciate the difficulties of working on a hazardous construction platform without actually being on-site. The repetitive "on-the-job" learning experience will narrow the experiential gap between a novice and an expert. Hence, the virtual models will create new intuitive understanding about the dangers involved in construction operations.
Heng *et al.* (2016, p.106)	"BBS has been shown to be an effective methodology for improving safety behavior in construction work, but has failed to be widely adopted because it is dependent on a manual and experienced inspection process, and lacks accurate assessment and timely feedback. This paper solves this problem by providing an effective approach to automatically identifying intrusion behaviors with positioning technology and assesses the personal safety performance of workers according their response to danger warnings".
Hou *et al.* (2017, p.39)	"AR/VR training system can enhance the quality of training to avoid errors and failures during the task. The outcomes of this study could help the AEC industry to significantly improve the productivity and safety of its workforce, while providing paradigms of transformative learning process of pedagogically adopting ICT in curricula and assessment regimes".
Hsiao *et al.* (2005, p.1239)	"The result, coupled with the improvement of walking instability score at the third trial, implied that individuals with limited experience of working at elevation may benefit from virtual-height training, which would help to develop adequate balance control skills, before they enter an elevated construction job in order to reduce their fall risks".

Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training (cont.).

Paper	Contribution
Juang *et al.* (2013, p.517)	"Developed training system improved the realism of a virtual training system and reduced the experiential difference between operating virtual and real cranes".
Khoury and Kamat (2009, p.487)	"In order to interactively display information retrieved using the described methodology to the mobile user, a new class of display and interaction devices is needed. It is important that any information support provided to the user must not come at a safety and efficiency cost".
Ki (2011, p.737)	"The simulator described in this article was designed to mimic the operation of real equipment, although there is no actual motion. The proposed simulator can be used not only for beginners' training, but also for safe skill improvement. The evaluation level can be adjusted to the trainee's level and can guide the trainee to the desired goal using evaluation logic and additional information, such as the distance between the wall and the basket".
Kim *et al.* (2017, p.401, p.402)	"The study proved that a vision-based wearable device with augmented reality could play a crucial role in the delivery of hazard information to construction workers. As the wearable device does not require any operational control, the construction workers can continually be fed the safety information without them having to stop their ongoing task.? "The hazard information is displayed as augmented reality in the wearable device so that a worker can effectively recognize dangerous situations without any behavioral constraints".
Le *et al.* (2015, p.19)	"Applying the framework of the system to continuous construction processes would assist learners in practicing virtual construction activities and in realizing the kinds of potential risks involved in real situations. Then, the accident on-site can be reduced proactively".
Lee *et al.* (2014, p.734)	"Based on this recognition, this study introduced a real-time location-based construction labor safety management system. For this purpose, a real-time locating system (RTLS) was applied to track location of workers on a real-time basis, and a location monitoring system was developed where workers' locations and potential dangerous areas are visualized. Finally, an alarming technology was utilized to warn endangered workers. Integrating these technologies, the developed system can assist the management and control of direct causes of accidents in real-time".
Li *et al.* (2012a, p.504)	"The use of VSAS can assess the users' knowledge in identifying; unsafe working environments; unsafe working attitudes; and unsafe working methods / procedures".

Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training (cont.).

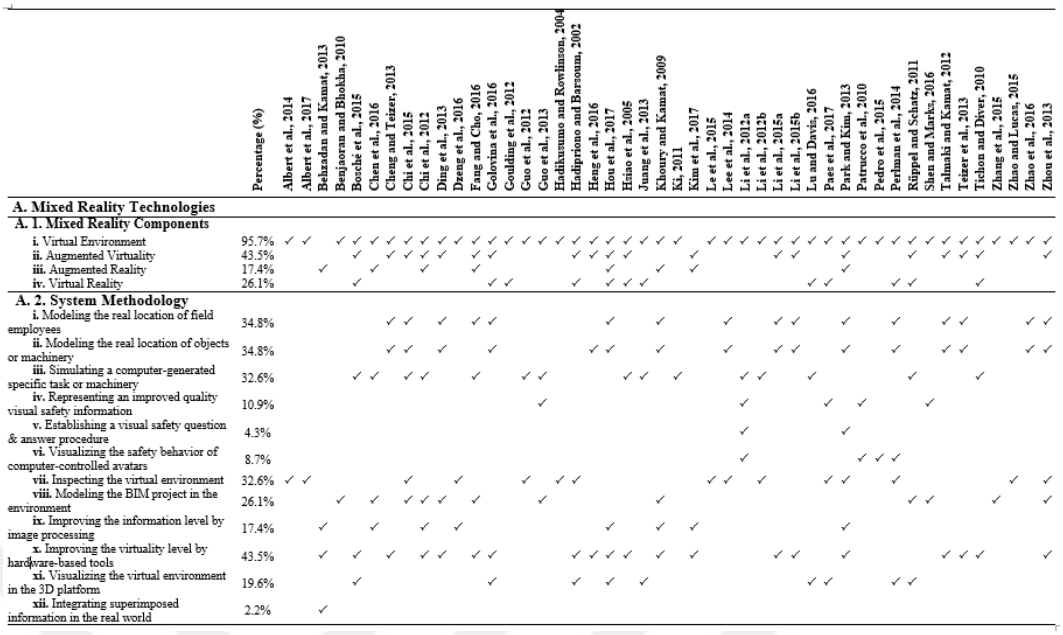
Paper	Contribution
Li *et al.* (2012b, p.646)	"Game engines and safety training provides a close-to-reality 3D environment in which trainees can learn and practice their knowledge. This system, termed the multiuser virtual safety training system, comprises four developed functions to allow trainees to learn comprehensive construction processes in a virtual and risk-free environment".
Li *et al.* (2015a, p.174)	"Thus, hazards were detected and timely proactive warnings provided to avoid accidents; training was visualized vividly and assessed quantitatively; real-time or post-event feedback helped workers change their risky behaviors and improve the training programme"
Li *et al.* (2015b, p.63, p.64)	"The proposed system improves construction safety by alerting workers to proximate construction hazards". "Thus, in the search for a suitable means of proactive safety management of construction work, the CSS system offers a clear practical advantage by providing a more flexible aid in the automatic positional monitoring of site workers".
Lu and Davis (2016, p.193)	"This study has shown that the addition of background sound to a high fidelity virtual reality simulator enhances people's sense of presence by creating a high level of normality".
Paes *et al.* (2017, p.301)	"The greater accuracy in identifying certain spatial aspects and distances using the immersive platform implies a better understanding of the spatial arrangement of the 3D model. In other words, if the user is able to, for instance, better estimate distances - when the virtual estimate is closer to the estimate in the physical environment - it means that he/she better understood the spatial arrangement of that virtual model".
Park and Kim (2013, p.102)	"In the previous visualization technology applications, it has been well proven that virtual 3D site models have a great potential to improve the identification of safety risks as well as worker's risk cognition and safety education. Meanwhile, 3D model based role-playing education immerses the worker in the situation. From the virtual reality, the worker could learn the exact risk in their job site".
Patrucco *et al.* (2010, p.515)	"According to the interviewed workers, formal training programs came out to fail in addressing the "real-life" hazards; they also emphasized that practical, hands-on knowledge and a focus on job-specific skill sets are critical to the effectiveness of the training".
Pedro *et al.* (2015, p.9) Perlman *et al.* (2014, p.30)	"Results show that the VSES has the potential to integrate safety with construction methods education, improve learners' safety knowledge, and hazard identification abilities. "Most subjects identified more hazards correctly in the virtual environment than did subjects who studied photographs and documents".

Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training (cont.).

Paper	Contribution
Rüppel and Schatz (2011, p.610)	"However, eventually, the system developed according to the approach described in this article, should enable new analysis methods for studying the human behavior in extreme situations within the performance-based fire safety design and it is assumed that particularly, the human factors regarding the interaction occupant-building structure could be explored in a new way. This could help to improve existing or develop new evacuation models".
Shen and Marks (2016)	"Contributions of this research include a near-miss visualization user interface allows construction personnel to view near misses throughout a construction project to identify hazardous areas and frequency of near misses as well as feasibility study data of the created tool".
Talmaki and Kamat (2012, p.15)	"The combined sensor input and user input to control a backhoe excavator model demonstrates the potential use of the HV framework to assist excavation crews to prevent buried utility strikes".
Teizer *et al.* (2013, p.67)	"The results, including worker feedback, indicate that both real-time location tracking and data visualization technologies possess advantages than can improve safety understanding".
Tichon and Diver (2010, p.229)	"As investigations into the application and potential of this technology in the construction sector expand, it may be that it also has a role to play in the improvement of existing plant operator techniques perhaps value adding to refresher training to address both the expansion and the improvement of operator skills sets and identifying operator bad habits which can contribute to accidents".
Zhang *et al.* (2015, p.29)	"The developed construction safety ontology enables more effective inquiry of safety knowledge, which is the first step towards automated safety planning for JHA using BIM"
Zhao and Lucas (2015, p.66)	"This interactive training allows for a development of cognitive abilities and awareness that typically improve the users' comprehension of training material. It may also contribute to build the safety culture which promotes construction workers' safety practical routines".

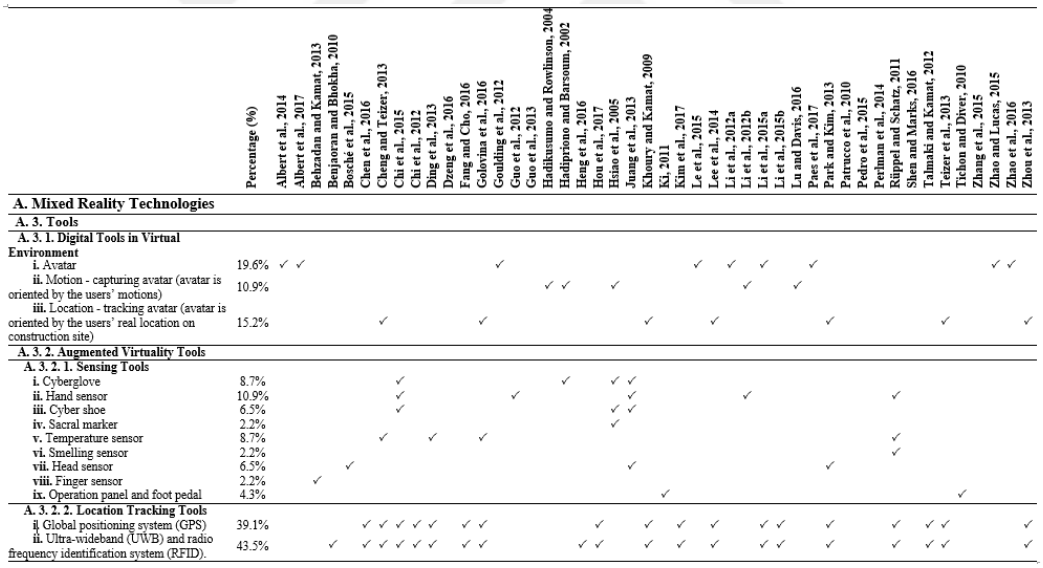
Table 2.3. The contribution of the studies in the context of mixed reality-based construction safety training (cont.).

Paper	Contribution
Zhao *et al.* (2016, p.806)	"Researchers presented an innovative approach to OSH risk mitigation that is integrated with safety culture. A "safety culture' might directly benefit workers' safety performance and also get forged into workers' safety practices through cutting-edge information technology with the concept of habitus. A prototype application for OSH training which used mobile virtual reality (MVR) technology was demonstrated to help establish the safety culture in workers' daily practices, and ultimately to mitigate the OSH risks in construction projects. At last, a preliminary validation test was conducted to examine the prototype's performance. The test results strongly supported human behavior influence and safe work knowledge comprehension by the application".
Zhou *et al.* (2013)	"The results show that the proposed approach can be a tool of collaboration, virtual analysis and prediction for designers, site project managers, safety engineers and other participants. With real-time safety status visualized, it can detect safety risks before and during the construction process and then provide preventive measures. Thus, timely decisions can be made to avoid accidents.?



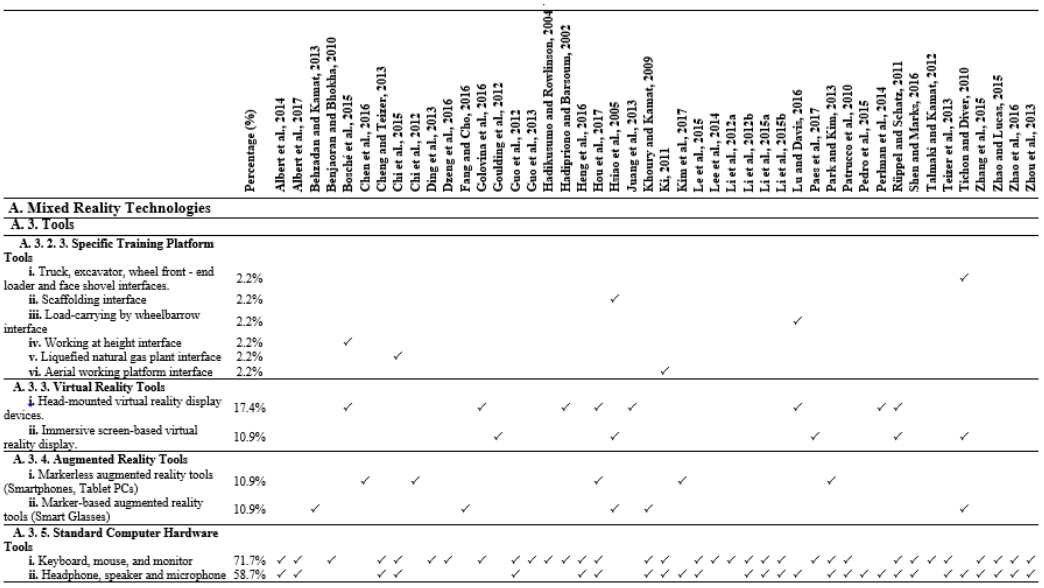
	Percentage (%)	Albert et al., 2014	Albert et al., 2017	Behzadan and Kamat, 2013	Benjaoran and Bhokha, 2010	Bosche et al., 2015	Chen et al., 2016	Cheng and Teizer, 2013	Chi et al., 2015	Chi et al., 2012	Ding et al., 2013	Dzeng et al., 2016	Fang and Cho, 2016	Golovina et al., 2016	Goulding et al., 2012	Guo et al., 2012	Guo et al., 2013	Hadikusumo and Rowlinson, 2004	Hadiprono and Barsoum, 2002	Heng et al., 2016	Hou et al., 2017	Hsiao et al., 2005	Juang et al., 2013	Khoury and Kamat, 2009	Ki, 2011	Kim et al., 2017	Le et al., 2015	Lee et al., 2014	Li et al., 2012a	Li et al., 2012b	Li et al., 2015a	Li et al., 2015b	Lu and Davis, 2016	Paes et al., 2017	Park and Kim, 2013	Patrucco et al., 2010	Pedro et al., 2015	Perlman et al., 2014	Rüppel and Schatz, 2011	Shen and Marks, 2016	Talmaki and Kamat, 2012	Teizer et al., 2013	Tichon and Diver, 2010	Zhang et al., 2015	Zhao and Lucas, 2015	Zhao et al., 2016	Zhou et al., 2013
A. Mixed Reality Technologies																																														
A. 1. Mixed Reality Components																																														
i. Virtual Environment	95.7%	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ii. Augmented Virtuality	43.5%					✓		✓	✓	✓	✓		✓	✓					✓	✓	✓	✓				✓					✓	✓		✓				✓		✓	✓	✓			✓
iii. Augmented Reality	17.4%			✓			✓			✓			✓								✓			✓		✓																				
iv. Virtual Reality	26.1%					✓								✓	✓				✓		✓	✓	✓															✓	✓				✓			
A. 2. System Methodology																																														
i. Modeling the real location of field employees	34.8%							✓	✓		✓		✓	✓							✓			✓			✓			✓	✓		✓			✓			✓	✓			✓	✓
ii. Modeling the real location of objects or machinery	34.8%							✓	✓		✓			✓						✓	✓			✓			✓			✓	✓		✓			✓			✓	✓			✓	✓
iii. Simulating a computer-generated specific task or machinery	32.6%					✓	✓		✓	✓			✓			✓	✓					✓	✓		✓			✓	✓			✓		✓				✓				✓			
iv. Representing an improved quality visual safety information	10.9%																✓											✓					✓		✓				✓							
v. Establishing a visual safety question & answer procedure	4.3%																											✓						✓											
vi. Visualizing the safety behavior of computer-controlled avatars	8.7%																											✓							✓	✓	✓								
vii. Inspecting the virtual environment	32.6%	✓	✓						✓			✓				✓		✓	✓								✓	✓	✓				✓	✓				✓					✓		✓
viii. Modeling the BIM project in the environment	26.1%				✓		✓		✓	✓	✓		✓				✓							✓														✓	✓			✓			✓
ix. Improving the information level by image processing	17.4%			✓			✓			✓		✓									✓			✓	✓									✓											
x. Improving the virtuality level by hardware-based tools	43.5%			✓		✓		✓		✓	✓		✓	✓					✓	✓	✓	✓		✓	✓					✓	✓				✓					✓	✓	✓			✓
xi. Visualizing the virtual environment in the 3D platform	19.6%					✓					✓				✓				✓	✓		✓									✓	✓					✓	✓						
xii. Integrating superimposed information in the real world	2.2%			✓																																									

Figure 2.2. The systematic review of the mixed reality technologies 1.



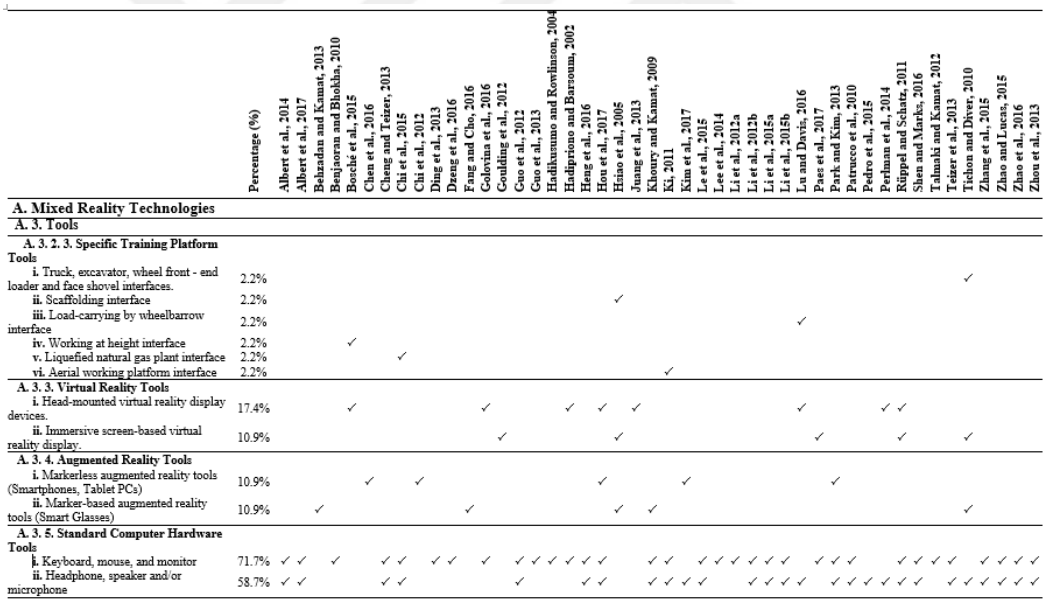
	Percentage (%)	Albert et al., 2014	Albert et al., 2017	Behzadan and Kamat, 2013	Benjaoran and Bhokha, 2010	Bosche et al., 2015	Chen et al., 2016	Cheng and Teizer, 2013	Chi et al., 2015	Chi et al., 2012	Ding et al., 2013	Dzeng et al., 2016	Fang and Cho, 2016	Golovina et al., 2016	Goulding et al., 2012	Guo et al., 2012	Guo et al., 2013	Hadikusumo and Rowlinson, 2004	Hadiprono and Barsoum, 2002	Heng et al., 2016	Hou et al., 2017	Hsiao et al., 2005	Juang et al., 2013	Khoury and Kamat, 2009	Ki, 2011	Kim et al., 2017	Le et al., 2015	Lee et al., 2014	Li et al., 2012a	Li et al., 2012b	Li et al., 2015a	Li et al., 2015b	Lu and Davis, 2016	Paes et al., 2017	Park and Kim, 2013	Patrucco et al., 2010	Pedro et al., 2015	Perlman et al., 2014	Rüppel and Schatz, 2011	Shen and Marks, 2016	Talmaki and Kamat, 2012	Teizer et al., 2013	Tichon and Diver, 2010	Zhang et al., 2015	Zhao and Lucas, 2015	Zhao et al., 2016	Zhou et al., 2013
A. Mixed Reality Technologies																																														
A. 3. Tools																																														
A. 3. 1. Digital Tools in Virtual Environment																																														
i. Avatar	19.6%	✓	✓											✓												✓		✓		✓		✓			✓									✓	✓		
ii. Motion - capturing avatar (avatar is oriented by the users' motions)	10.9%																	✓	✓		✓								✓			✓														
iii. Location - tracking avatar (avatar is oriented by the users' real location on construction site)	15.2%							✓					✓											✓			✓							✓							✓				✓
A. 3. 2. Augmented Virtuality Tools																																													
A. 3. 2. 1. Sensing Tools																																														
i. Cyberglove	8.7%								✓										✓			✓	✓																							
ii. Hand sensor	10.9%								✓							✓							✓						✓										✓						
iii. Cyber shoe	6.5%								✓													✓	✓																						
iv. Sacral marker	2.2%																					✓																								
v. Temperature sensor	8.7%							✓			✓		✓																										✓							
vi. Smelling sensor	2.2%																																						✓							
vii. Head sensor	6.5%				✓																		✓											✓											
viii. Finger sensor	2.2%			✓																																									
ix. Operation panel and foot pedal	4.3%																							✓																			✓			
A. 3. 2. 2. Location Tracking Tools																																														
i. Global positioning system (GPS)	39.1%						✓	✓	✓	✓	✓		✓	✓							✓			✓	✓	✓		✓		✓	✓			✓				✓		✓	✓				✓
ii. Ultra-wideband (UWB) and radio frequency identification system (RFID).	43.5%				✓		✓	✓	✓	✓	✓		✓	✓						✓	✓			✓	✓	✓		✓	✓			✓				✓		✓	✓				✓

Figure 2.3. The systematic review of the mixed reality technologies 2.



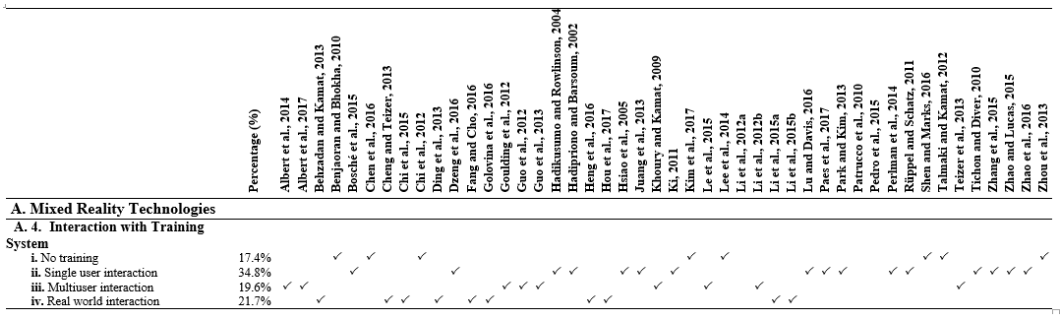
	Percentage (%)	Albert et al., 2014	Albert et al., 2017	Behzadan and Kamat, 2013	Benjaoran and Bhokha, 2010	Bosché et al., 2015	Chen et al., 2016	Cheng and Teizer, 2013	Chi et al., 2015	Chi et al., 2012	Ding et al., 2013	Dzeng et al., 2016	Fang and Cho, 2016	Golovina et al., 2016	Goulding et al., 2012	Guo et al., 2012	Guo et al., 2013	Hadikusumo and Rowlinson, 2004	Hadipriono and Barsoum, 2002	Heng et al., 2016	Hou et al., 2017	Hsiao et al., 2005	Juang et al., 2013	Khoury and Kamat, 2009	Ki, 2011	Kim et al., 2017	Le et al., 2015	Lee et al., 2014	Li et al., 2012a	Li et al., 2012b	Li et al., 2015a	Li et al., 2015b	Lu and Davis, 2016	Paes et al., 2017	Park and Kim, 2013	Patrucco et al., 2010	Pedro et al., 2015	Perlman et al., 2014	Rüppel and Schatz, 2011	Shen and Marks, 2016	Talmaki and Kamat, 2012	Teizer et al., 2013	Tichon and Diver, 2010	Zhang et al., 2015	Zhao and Lucas, 2015	Zhao et al., 2016	Zhou et al., 2013
A. Mixed Reality Technologies																																														
A. 3. Tools																																														
A. 3. 2. 3. Specific Training Platform Tools																																														
i. Truck, excavator, wheel front - end loader and face shovel interfaces.	2.2%																																									✓			
ii. Scaffolding interface	2.2%																					✓																							
iii. Load-carrying by wheelbarrow interface	2.2%																															✓													
iv. Working at height interface	2.2%					✓																																							
v. Liquefied natural gas plant interface	2.2%								✓																																					
vi. Aerial working platform interface	2.2%																								✓																					
A. 3. 3. Virtual Reality Tools																																													
i. Head-mounted virtual reality display devices.	17.4%						✓							✓						✓		✓		✓									✓						✓	✓						
ii. Immersive screen-based virtual reality display.	10.9%														✓							✓											✓									✓				
A. 3. 4. Augmented Reality Tools																																													
i. Markerless augmented reality tools (Smartphones, Tablet PCs)	10.9%						✓		✓												✓				✓										✓										
ii. Marker-based augmented reality tools (Smart Glasses)	10.9%			✓									✓									✓	✓																				✓			
A. 3. 5. Standard Computer Hardware Tools																																													
i. Keyboard, mouse, and monitor	71.7%	✓	✓		✓			✓	✓			✓	✓		✓		✓	✓	✓	✓	✓			✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓		✓	✓	✓	✓
ii. Headphone, speaker and microphone	58.7%	✓	✓					✓	✓							✓					✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓

Figure 2.4. The systematic review of the mixed reality technologies 3.



	Percentage (%)	Albert et al., 2014	Albert et al., 2017	Behzadan and Kamat, 2013	Benjaoran and Bhokha, 2010	Bosché et al., 2015	Chen et al., 2016	Cheng and Teizer, 2013	Chi et al., 2015	Chi et al., 2012	Ding et al., 2013	Dzeng et al., 2016	Fang and Cho, 2016	Golovina et al., 2016	Goulding et al., 2012	Guo et al., 2012	Guo et al., 2013	Hadikusumo and Rowlinson, 2004	Hadipriono and Barsoum, 2002	Heng et al., 2016	Hou et al., 2017	Hsiao et al., 2005	Juang et al., 2013	Khoury and Kamat, 2009	Ki, 2011	Kim et al., 2017	Le et al., 2015	Lee et al., 2014	Li et al., 2012a	Li et al., 2012b	Li et al., 2015a	Li et al., 2015b	Lu and Davis, 2016	Paes et al., 2017	Park and Kim, 2013	Patrucco et al., 2010	Pedro et al., 2015	Perlman et al., 2014	Rüppel and Schatz, 2011	Shen and Marks, 2016	Talmaki and Kamat, 2012	Teizer et al., 2013	Tichon and Diver, 2010	Zhang et al., 2015	Zhao and Lucas, 2015	Zhao et al., 2016	Zhou et al., 2013
A. Mixed Reality Technologies																																														
A. 3. Tools																																														
A. 3. 2. 3. Specific Training Platform Tools																																														
i. Truck, excavator, wheel front - end loader and face shovel interfaces.	2.2%																																								✓			
ii. Scaffolding interface	2.2%																					✓																							
iii. Load-carrying by wheelbarrow interface	2.2%																															✓													
iv. Working at height interface	2.2%					✓																																							
v. Liquefied natural gas plant interface	2.2%								✓																																					
vi. Aerial working platform interface	2.2%																								✓																					
A. 3. 3. Virtual Reality Tools																																													
i. Head-mounted virtual reality display devices.	17.4%						✓							✓						✓		✓		✓									✓						✓	✓						
ii. Immersive screen-based virtual reality display.	10.9%														✓							✓											✓									✓				
A. 3. 4. Augmented Reality Tools																																													
i. Markerless augmented reality tools (Smartphones, Tablet PCs)	10.9%						✓		✓												✓				✓										✓										
ii. Marker-based augmented reality tools (Smart Glasses)	10.9%			✓									✓									✓	✓																				✓			
A. 3. 5. Standard Computer Hardware Tools																																													
i. Keyboard, mouse, and monitor	71.7%	✓	✓		✓			✓	✓			✓	✓		✓		✓	✓	✓	✓	✓			✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓		✓	✓	✓	✓
ii. Headphone, speaker and/or microphone	58.7%	✓	✓					✓	✓							✓					✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓

Figure 2.5. The systematic review of the mixed reality technologies 4.



	Percentage (%)	Albert et al., 2014	Albert et al., 2017	Behzadan and Kamat, 2013	Benjaoran and Bhokha, 2010	Bosché et al., 2015	Chen et al., 2016	Cheng and Teizer, 2013	Chi et al., 2015	Chi et al., 2012	Ding et al., 2013	Dzeng et al., 2016	Fang and Cho, 2016	Golovina et al., 2016	Goulding et al., 2012	Guo et al., 2012	Guo et al., 2013	Hadikusumo and Rowlinson, 2004	Hadipriono and Barsoum, 2002	Heng et al., 2016	Hou et al., 2017	Hsiao et al., 2005	Juang et al., 2013	Khoury and Kamat, 2009	Ki, 2011	Kim et al., 2017	Le et al., 2015	Lee et al., 2014	Li et al., 2012a	Li et al., 2012b	Li et al., 2015a	Li et al., 2015b	Lu and Davis, 2016	Paes et al., 2017	Park and Kim, 2013	Patrucco et al., 2010	Pedro et al., 2015	Perlman et al., 2014	Rüppel and Schatz, 2011	Shen and Marks, 2016	Talmaki and Kamat, 2012	Teizer et al., 2013	Tichon and Diver, 2010	Zhang et al., 2015	Zhao and Lucas, 2015	Zhao et al., 2016	Zhou et al., 2013
A. Mixed Reality Technologies																																														
A. 4. Interaction with Training System																																														
i. No training	17.4%				✓		✓			✓										✓	✓				✓		✓		✓											✓	✓						✓
ii. Single user interaction	34.8%					✓						✓										✓	✓		✓							✓	✓	✓				✓	✓			✓	✓	✓	✓	
iii. Multiuser interaction	19.6%	✓	✓											✓	✓	✓	✓							✓			✓			✓												✓					
iv. Real world interaction	21.7%			✓				✓	✓		✓		✓	✓						✓	✓										✓	✓														

Figure 2.6. The systematic review of the mixed reality technologies 5.

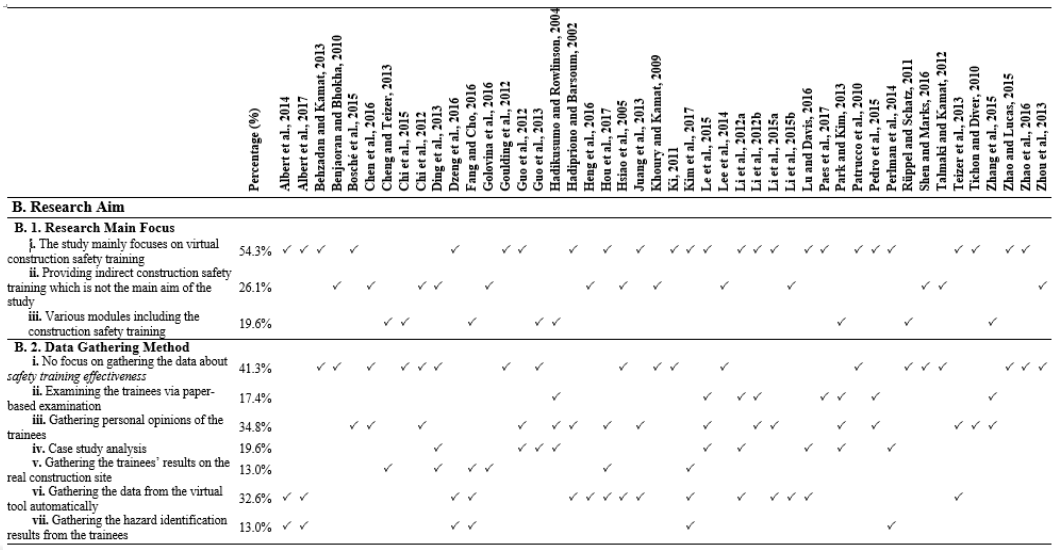


Figure 2.7. The systematic review of the research orientation 1.

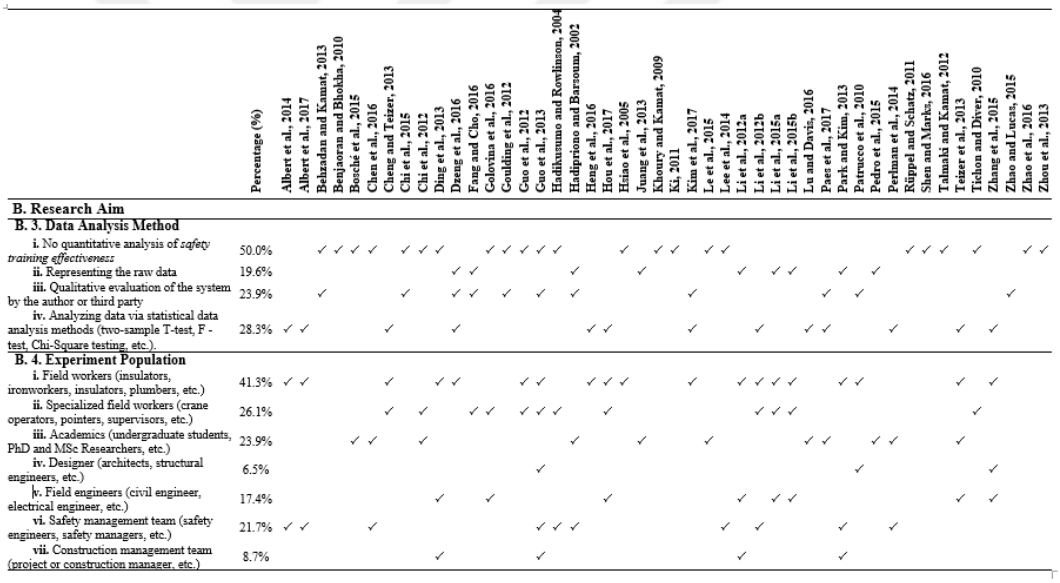
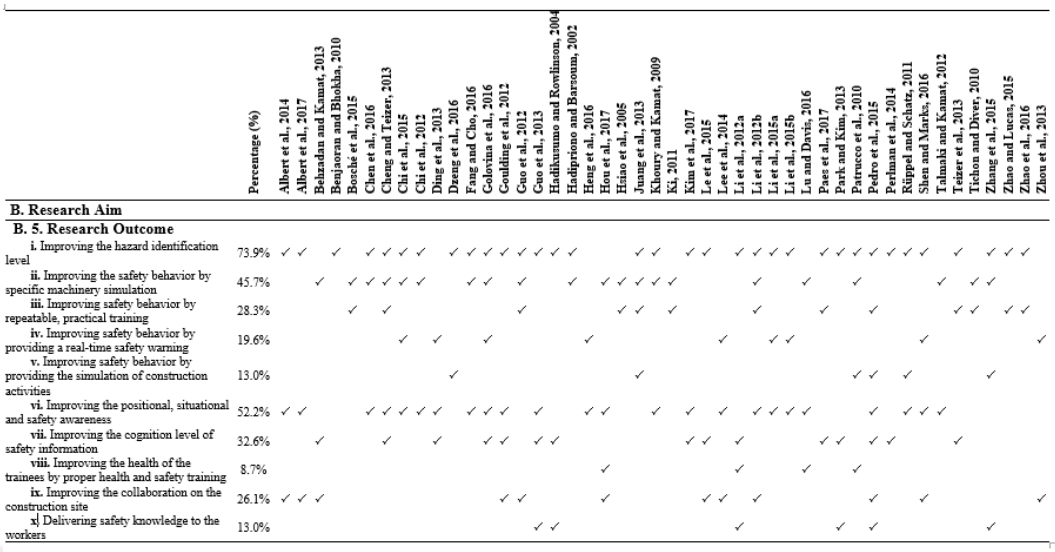


Figure 2.8. The systematic review of the research orientation 2.



	Percentage (%)	Albert et al., 2014	Albert et al., 2017	Behzadan and Kamat, 2013	Benjaoran and Bhokha, 2010	Bosché et al., 2015	Chen et al., 2016	Cheng and Teizer, 2013	Chi et al., 2015	Chi et al., 2012	Ding et al., 2013	Dzeng et al., 2016	Fang and Cho, 2016	Golovina et al., 2016	Goulding et al., 2012	Guo et al., 2012	Guo et al., 2013	Hadikusumo and Rowlinson, 2004	Hadipriono and Barsoum, 2002	Heng et al., 2016	Hou et al., 2017	Hsiao et al., 2005	Juang et al., 2013	Khoury and Kamat, 2009	Ki, 2011	Kim et al., 2017	Le et al., 2015	Lee et al., 2014	Li et al., 2012a	Li et al., 2012b	Li et al., 2015a	Li et al., 2015b	Lu and Davis, 2016	Paes et al., 2017	Park and Kim, 2013	Patrucco et al., 2010	Pedro et al., 2015	Perlman et al., 2014	Rüppel and Schatz, 2011	Shen and Marks, 2016	Talmaki and Kamat, 2012	Teizer et al., 2013	Tichon and Diver, 2010	Zhang et al., 2015	Zhao and Lucas, 2015	Zhao et al., 2016	Zhou et al., 2013
B. Research Aim																																													
B. 5. Research Outcome																																													
i. Improving the hazard identification level	73.9%	✓	✓		✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓				✓	✓		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	
ii. Improving the safety behavior by specific machinery simulation	45.7%			✓		✓	✓	✓	✓	✓			✓	✓		✓			✓	✓	✓	✓	✓	✓				✓			✓			✓					✓		✓	✓			
iii. Improving safety behavior by repeatable, practical training	28.3%					✓		✓								✓					✓	✓		✓				✓			✓		✓		✓						✓	✓		✓	✓	
iv. Improving safety behavior by providing a real-time safety warning	19.6%								✓		✓			✓					✓								✓			✓	✓								✓						✓
v. Improving safety behavior by providing the simulation of construction activities	13.0%											✓											✓												✓	✓		✓				✓			
vi. Improving the positional, situational and safety awareness	52.2%	✓	✓				✓	✓	✓	✓	✓		✓	✓	✓		✓			✓	✓			✓		✓	✓		✓	✓	✓	✓	✓			✓		✓	✓	✓					
vii. Improving the cognition level of safety information	32.6%			✓			✓			✓			✓	✓		✓	✓								✓	✓		✓					✓	✓		✓	✓				✓				
viii. Improving the health of the trainees by proper health and safety training	8.7%																			✓								✓			✓			✓										
ix. Improving the collaboration on the construction site	26.1%	✓	✓	✓											✓	✓				✓					✓	✓		✓							✓			✓						✓
x	Delivering safety knowledge to the workers	13.0%															✓	✓										✓						✓	✓						✓			

Figure 2.9. The systematic review of the research orientation 3.

2.5.1. Virtual Environment

In the systematic review, it was noted that around 95% of the reviewed studies used the virtual environment in their safety training procedure (see Figure 2.2, A. 1. i.). One of the purposes of using virtual environments, which is the most common utilized tool in the safety training process (see Figure 2.2, A. 1. i.), is to model the real location of the field employees and objects on the construction site by using sensing tools such as GPS, UWB, and RFID (see Figure 2.2, A. 2. i. and ii.; Figure 2.2, A. 3. 2. 2. i. and ii.). In terms of modeling the real location of objects and field employees, several studies (e.g. Ding *et al.*, 2013; Chi *et al.*, 2015) develop real-time early safety warning systems to the workers, thus, they improve their safety awareness (see Figure 2.5, B. 2. iv. and vi.). In terms of modeling the locations on the virtual construction site, location tracking-based environments are not only useful for improving situational awareness, but they also provide an adequate safety monitoring process for safety managers.

Alternatively, around 30% of all the studies have focused on inspecting the virtual environment (see Figure 2.4, A. 2. vii.). In an inspection-based virtual environment, the trainees could walk through with their avatars in the virtual terrain, and they could interact with the virtual objects and other trainees. The review of the inspection-

based virtual environment studies, (e.g. Albert *et al.*, 2014; Li *et al.*, 2012a; Li *et al.*, 2012b; Le *et al.*, 2015; Guo *et al.*, 2012; Zhao and Lucas, 2015) reveals that the method is quite suitable to improve the hazard identification level of the trainees (see Figure 2.5. B. 2. i.). Almost all of the studies that use inspection-based virtual environment claim that virtual environments improve the hazard identification level. Quite consistently, some researchers (e.g. Albert *et al.*, 2014; Park and Kim, 2013) have reported that adequate hazard identification should involve the mental visualization of construction tasks. In this sense, since the inspection-based virtual environments enable the display of construction tasks, the mental visualization level of the trainees are also improved (Albert *et al.*, 2014). As an alternative procedure, few studies (e.g. Li *et al.*, 2012; Park and Kim, 2013) used the inspection-based environment to develop a question and answer procedure by the integration of questions with the relevant risk source involved in the virtual media. In this procedure, when a trainee moves closer to virtual construction equipment, the specific safety question appears, and trainees could answer the question by clicking the correct response in the virtual media. The main aim of this procedure is to prevent the low comprehension of safety questions in the paper-based examination methodology. In other words, these studies (e.g. Li *et al.*, 2012; Park and Kim, 2013) argue that answering a question in a 3D environment is more effective than paper-based methods, in terms of understanding the question properly thus improving cognitive level. Also, this technique significantly improves the material-based risk source categorization. As a similar approach to question and answer procedure, several studies (e.g. Patrucco *et al.*, 2010; Perlman *et al.*, 2014; Pedro *et al.*, 2015) have focused on visualizing the safety behavior of non-playable characters (NPCs) (see Figure 2.4. A. 2. vi). These studies have used NPCs to represent the correct actions during construction tasks. The primary goal of visualizing the functions of daily construction tasks is to enhance the understandability level of safety information, since, the environment enables monitoring the situations from different perspectives in the virtual space. Even though using NPCs and the question and answer procedure serve the same purpose, the learning and evaluation methods used in these approaches are different. In other words, the question and answer method mainly focuses on evaluating the safety information level of the trainees through an alternative information delivery method by using virtual models. Using NPCs aims to enhance the cognitive level of

the trainees through visualizing correct safety behavior.

As another purpose, some studies (e.g., Guo *et al.*, 2013; Hadikusumo and Rowlinson, 2004) used a BIM-based environment to represent the construction tasks that involve safety hazards. In general, the BIM-based environment focuses on simulating a construction task in a 3D environment (e.g., Guo *et al.*, 2013; Hadikusumo and Rowlinson, 2004) and these studies aim to understand the potential outcomes of a specific task. In this sense, target populations such as designers, crane operators could improve their knowledge about the design and simulated tasks; therefore, their safety knowledge level indirectly improves. Moreover, these studies (e.g. Guo *et al.*, 2013; Hadikusumo and Rowlinson, 2004) used a BIM-based environment to capture and transfer the safety knowledge. In general, safety knowledge is usually transferred to the workers via conventional methods such as 2D architectural drawings, class lectures, videos, and written materials. However, these methods fail to consider the third dimension, and therefore the cognitive level of workers is quite limited due to the lack of spatial awareness. In this sense, many researchers (e.g. Guo *et al.*, 2013; Hadikusumo and Rowlinson, 2004) aim to fill this gap by using a BIM-based environment, since visualized safety information about a specific task in a 3D environment provides the third dimension representation. As a result, the cognitive level of workers about the provided safety information could be improved by the BIM-based environment visualization. On the other hand, these studies do not primarily aim to enhance the safety training process. In the review, our analysis shows that more than half of the reviewed studies do not mainly focus on improving safety training process and many of these studies use BIM-based environment. In this sense, BIM-based environment has a high potential to impact the safety training process directly.

In the lights of this research, it is concluded that virtual environments have significant strengths in the construction safety training field. First of all, virtual environments provide repeatable hands-on practice. Thus, trainees could reinforce their safety knowledge by repetitively practicing in a risk-free environment. Besides, trainees can also familiarize themselves with the equipment they work and the inherent hazards of the tasks by repeatable hands-on practice. Secondly, virtual environments enhance

the mental visualization of the trainees by taking into account the third dimension; in this way, trainees could identify on-site hazards easily (Albert *et al.*, 2014). Moreover, Albert and colleagues (2014) state that the effectiveness of the virtual environment-based safety training is directly related to the visual quality of the virtual media. In other words, inadequate visual quality could lead to inefficiency in safety training. In this sense, the developers should ensure visualization quality.

2.5.2. Virtual Reality and Augmented Virtuality

As another tool in reality-virtuality continuum, virtual reality is an effective form of display settings of the virtual environments. Shortly, the virtual reality system is based on the conversion of the virtual environment into a surrounding screen or head-mounted display, which improves the visualization quality of the media (Squelch, 2001). Moreover, virtual reality display sets improve the visualization quality of the virtual environment by integrating third dimension (Milgram and Kishino, 1994). In addition to the improvement of visualization quality, virtual reality display sets also improve the interaction quality between user and virtual environment. By utilizing head pose to approximate a users' line-of-sight for real-time image rendering and interaction in the 3D visualization applications, users could interact with the environment by head movement, which is similar to the real-life (Hua *et al.*, 2006). In other words, orienting the avatars by using head movement could be more realistic, compared to the orientation by keyboard and mouse layout. In this sense, virtual reality display sets can provide one of the significant features of augmented virtuality as improving interaction quality. Several studies (e.g. Golovina *et al.*, 2016; Hou *et al.*, 2017; Bosché *et al.*, 2016; Hadipriono and Barsoum; 2002) used immersive virtual environments, to provide better visualization. In these studies, head-mounted or immersive screen-based virtual reality display sets have improved the visualization quality. In general, immersive virtual environments are a standard procedure to develop a simulation platform for a specific machinery (e.g. Patrucco *et al.*, 2010; Guo *et al.*, 2013; Bosché *et al.*, 2016; Li *et al.*, 2012a; Guo *et al.*, 2012; Tichon and Diver, 2010; Hsiao *et al.*, 2005). The main objective of these studies is to provide hands-on training to convey the safety information through practicing in a similar environment to the specific construction

task but in a risk-free virtual environment.

Another component to improve the interaction quality is augmented virtuality tools. The analysis shows that around 44% of the studies used augmented virtuality (see Figure 2.4, A. 1. ii.). The augmented virtuality technologies improve the interaction quality by integrating proper tools such as cyber gloves, hand sensors, cyber shoes, etc. into the simulation. Commonly, a limited number of studies (e.g. Chi *et al.*, 2015; Tichon and Diver, 2010; Hsiao *et al.*, 2005) report that some construction tasks such as crane operations and scaffolding require on-site proactive action, to effectively prevent the hazards. These studies (e.g. Chi *et al.*, 2015; Tichon and Diver, 2010; Hsiao *et al.*, 2005) have used specific training platforms, to simulate complex construction operations. Thus, they could be able to improve the safety behavior of trainees before interacting with the real systems on construction sites (Li *et al.*, 2012a).

2.5.3. Augmented Reality

Augmented reality is another major component of mixed reality that is used in around 22% of the studies (see Figure 2.4. A. 3. 4.). In general augmented reality has been used, to integrate superimposed information in the real world and to improve the safety information level by image processing (Wang and Dunston, 2005). Superimposed information and image processing methods provide enhanced visual information for the construction equipment in the real world. Thus, trainees could improve their safety knowledge by immediately gathering relevant safety information about the construction equipment and tasks. Using augmented reality technologies provides significant benefits for safety training effectiveness. Firstly, as mentioned, augmented reality enables trainees to reinforce their safety knowledge about real construction tasks by providing relevant information. Besides, augmented reality technologies establish a base for the trainees by improving mental cognition. For instance, Kim and colleagues (2016) developed an image-based safety assessment system by using augmented reality head-mounted display settings. Thus, Kim and colleagues aim to support both managers and workers to recognize hazardous situations. The study results show that a head-mounted augmented reality device successfully integrates the relevant safety

information to the real objects on the construction site and transfers this information to the construction workers and managers. In brief, the study used augmented reality to match the relevant information within the construction equipment by using image processing. Consistently, Park and Kim (2013) utilized mobile augmented reality software for improving the workers' cognition of the hazardous site elements; thus, a field inspection process was preferred. Case study analysis results show that workers and safety managers could be able to comprehend relevant site information more effectively by visualizing the exact environment of the construction field. Alternatively, Behzadan and Kamat (2013) developed a discovery-based learning platform by using augmented reality technology. The study results indicate that augmented reality-based safety training procedure provides collaborative learning, which is one of the crucial aspects of effective safety training. In their study, Behzadan and Kamat (2013) used telepresented augmented reality, to develop a hologram of a crane interface for managing the virtual objects. In other words, Behzadan and Kamat (2013) used augmented reality to interact with the virtual media, which is a common practice in augmented virtuality tools. Therefore, the significant aspect of effective safety training via augmented reality could be considered as clearly understanding the essential instructions about the construction tasks. In general, it is noticed that augmented reality have a high impact on construction safety. However, there are several limitations in these studies. Firstly, many of the studies that used augmented reality mainly focus on providing relevant information about the specified construction equipment. To gather data from the real construction site, augmented reality data receivers (e.g., mobile phones, cameras, etc.) continuously capture images from the viewpoint of trainees. For instance, when trainees direct a receiver to a scaffolding platform, it captures the image, analyzes the visual data and detects the equipment on the construction floor. However, construction activities are quite dynamic processes. In this context, a relevant data analysis algorithm based on the order of data retrieval may be useful. For instance, an augmented reality data algorithm that senses the unloading process of a crane movement could be sufficient. Secondly, during the systematic review, we also noticed that previous studies (e.g. Kim *et al.*, 2016; Park and Kim, 2013,) merely focused on testing the usability of the augmented reality tool in terms of matching the relevant information within the specified construction equipment. Therefore, these studies (e.g. Kim *et al.*,

2016; Park and Kim, 2013) aim to improve the hazard recognition level of the trainees for enabling them to take sudden action against hazards. In this context, analyzing the comprehension duration of the trainees could be quite suitable for further research. One can conclude that if the comprehension of the virtual object or information generated by augmented reality takes longer, the trainees will be distracted in the real world. As a result, a potential accident could occur on the actual construction site. For instance, if a tower crane operator concentrates on any information provided by the augmented reality for a long time during the loading activity, the operator might be distracted, and this can cause a potential on-site accident.

2.6. Measuring the Effectiveness of Safety Training

It was noted that previous studies utilized conventional data gathering methods such as collecting user feedbacks via interviews, conducting paper-based examinations and closed-ended questions or checking trainees' number of identified hazards for evaluating the effectiveness of developed virtual safety training method. (see Figure 2.5 - B.2.). Personal user feedback, which reflects the opinions of the target population, is usually collected through filling in a questionnaire. Although this method is suitable to gather the personal opinion of trainees, it is not appropriate to determine whether the safety training method is effective. Alternatively, several studies (e.g. Li *et al.*, 2012; Li *et al.*, 2015a; Le *et al.*, 2015; Guo *et al.*, 2012) used the semi-structured or structured interviews to collect feedback from the study groups. Even though these methods are helpful to gather personal opinion, they do not involve any quantitative data analysis, similarly to paper-based questionnaires. To quantitatively analyze the effectiveness of safety training procedure, eight studies (see Figure 2.5, B. 3. ii) provide a quantitative analysis by using the close-ended questions-based examination. Close-ended questions-based examination method provides a quantitative test method to avoid subjective evaluation. This method could be quite suitable to test the safety information level of trainees. However, close-ended questions fail to evaluate the safety knowledge and safety behavior of participants. In addition, some studies (see Figure 2.5, B. 2. vi) analyzed the effectiveness of developed method through the trainees' hazard identification level. The evaluation through the level of hazard identification could

be considered as a suitable method to evaluate the impact of method on trainees' risk recognition capabilities. However, this method could not be seen as a precise method, since, the hazard recognition has indirect impact of trainees' attitudes (Albert *et al.*, 2014). So, merely evaluating the hazard identification level of trainees do not prove that all trainees safety behavior is appropriate. Previous literature (Burke *et al.*, 2006; Albert *et al.*, 2014; Shapira and Elbaz; 2014; Bahn and Barratt-Pugh, 2014) highlight that an effective safety training method should involve knowledge development, safety awareness, behavior modeling, which could be only evaluated through safety behavior and safety attitudes of trainees. In this sense, automated analysis methods could provide an essential solution to bridge this gap. Several studies (e.g. Albert *et al.*, 2014; Patrucco *et al.*, 2010; Chen *et al.*, 2016; Cheng and Teizer; 2013; Hou *et al.*, 2017; Heng *et al.*, 2016; Lu *et al.*, 2016) used automated performance analysis methods to evaluate the effectiveness of developed virtual construction safety training tool. In the automated performance analysis, the safety performance of the target population is analyzed by evaluating the safety behavior of participants in the virtual environment. For instance, when a trainee misbehaves by walking under a load in the virtual environment, automated performance analysis method detects the error of the trainee simultaneously. A considerable amount of literature (i.e. Albert *et al.*, 2014; Patrucco *et al.*, 2010; Perlman *et al.*, 2014; Cheng and Teizer, 2013; Li *et al.*, 2015a; Tichon and Diver, 2010; Hsiao *et al.*, 2005). Hence, these studies (i.e. Albert *et al.*, 2014; Patrucco *et al.*, 2010; Perlman *et al.*, 2014; Cheng and Teizer, 2013; Li *et al.*, 2012b; Li *et al.*, 2015a; Tichon and Diver, 2010; Hsiao *et al.*, 2005) highlight that automated performance analysis methods are quite suitable to evaluate major parameters of the effective safety training such as collaboration, safety awareness, and safety knowledge by monitoring safety behavior of trainees in the virtual environment. In addition, the method is quite suitable to provide specific feedback. For this reason, automated performance analysis method could be considered as a more suitable approach than close-ended questions to analyze the actual safety performance of trainees in terms of subjective evaluation of safety behavior of trainee. Despite many advantages, automated performance analysis methods also contain several weaknesses. Automated performance analysis methods only examine the safety performance of trainees based on their safety behavior in the virtual environment. Therefore, this method fails to

examine real-life factors. For instance, psychological factors such as attention level, mental workload and emotional response of participants cannot be measured by automated performance analysis. Automated performance analysis methods have an algorithm that only recognizes and reports incorrect behavior in the virtual environment. Although the recognition is quite significant to represent user-specific feedback, the method cannot perceive the reasons for the errors such as not taking into account human-based psychological effects. Returning to the previous example, the automated performance analysis method can detect that a trainee is incorrectly passing under the load. However, the method cannot perceive the reason of this case such as participant might have a potential distraction at that moment, or would not see the load at that moment. For this reason, although automated performance analysis methods can examine behaviors to a certain extent, there are still shortcomings that are quite important. A more efficient method that allows evaluating the psychological effects that affect the safety behaviors of the trainees during virtual experiments is required. To bridge this gap, using eye-tracking tools could be suitable to evaluate safety behavior of the trainees in terms of psychological features such as attention level, mental workload, emotional response, and situational awareness.

2.6.1. Eye Tracking

According to Poole and Ball (2006:1), eye tracking is 'a technique whereby an individuals' eye movements are measured so that the researcher knows both where a person is looking at any given time and the sequence in which their eyes are shifting from one location to another. Therefore, using eye-tracking technologies could be suitable to provide an objective analysis of the behavior patterns of the trainees. Just and Carpenter (1976) developed the "eye-mind' hypothesis, which states whenever participants look at a virtual item such as an object or a word, at the same time they start to think about it. Moreover, as long as the staring continues, the users continue to think of the object. By proposing the eye-mind hypothesis, Just and Carpenter (1976) examined the correlation between eye movement and human responses such as attention and cognition, etc. A considerable amount of psychological and neuropsychological studies have stated that there is a significant relationship between attention and eye

movements (Yarbus 1967; Sun *et al.* 2008). In this sense, when a person often looks directly when they are attending to their stimuli (Duchowski 2007). In the construction safety training context, when a trainee visually interacts with a hazard source, it is quite likely for the trainee for identifying and perceiving the hazard and take safety actions (Hasanzadeh *et al.*, 2017). Thus, evaluating the safety behavior of the trainees becomes possible.

Eye-tracking technologies were used in various fields in the literature such as aviation (Anders, 2001), driving safety (Horng *et al.*, 2004), inspection (Duchowski *et al.*, 2001), linguistics (Andersson, 2006), web design (Eraslan *et al.*, 2016), marketing (Eraslan *et al.*, 2016), neuroscience (Khushaba *et al.*, 2013) and psychology (Isaacowitz, *et al.*, 2006; Levene, 1961). In the construction safety training context, only a few studies (e.g., Hasanzadeh *et al.*, 2016; Dzeng *et al.*, 2016; Bhoir *et al.*, 2015; Dzeng *et al.*, 2016; Pinheiro *et al.*, 2016) focused on eye-tracking technologies. Moreover, it has been recognized that the publications are quite recent. In this sense, although other sectors have been used, the eye-tracking methods for quite a long period, the construction industry drops behind all these fields.

In their study, Hasanzadeh and colleagues (2016) used mobile eye-tracker devices to quantify the situational awareness of the construction field workers. Study results show that situational awareness could be successfully evaluated via using mobile eye trackers. Similarly, Dzeng and colleagues (2016) compared the hazard identification level of non-experienced and experienced workers. Study results show that the experienced workers have been more successful compared to inexperienced workers in identifying on-site hazards. The primary importance of the study could be considered as successfully analyzing the visual data via fixations and the scan paths during the hazard identification. Quite consistent with the study conducted by Dzeng and colleagues (2016), Habibnezhad and colleagues (2016) also tested the potential correlation between the visual search strategies and hazard identification. Study results show that risk perception level affects the visual search strategies of the workers. Similarly, Bhoir and colleagues (2015) analyzed the attention level when a hazardous situation occurs on the construction site. Analysis results show that the attention level of the trainees

significantly improves when the workers see a high-risk situation. In another study, Pinheiro and colleagues (2016) analyzed the gazing behavior based on the original and virtual construction scenes. By this method, Pinheiro and colleagues objected to examine the impact of the gazing behavior on the hazard identification level. Analysis results show that the gazing behavior of the trainees is significantly different in terms of the hazard identification level of the users. Therefore, eye tracking technologies have high potential to evaluate safety training effectiveness, via integrated innovative algorithms to the safety behavior analysis.

3. RESEARCH METHODOLOGY

This chapter describes the research methodology adopted in this research. In the first part of the research methodology, according to the detailed literature review about safety training effectiveness parameters, several research hypotheses and statements are constructed. In the second part of the chapter, various tools and parameters such as V-SAFE, eye tracking devices are described, which are used during the experiments. Finally, data analysis methods and experimental procedure of the experiments are illustrated.

3.1. Hypotheses and Statements of the Research

3.1.1. Statement and Hypotheses Related to Information Transfer

To clarify the effectiveness of the safety training methods based on information transfer, the previous literature about information quality is reviewed. Information quality is defined by Huang and colleagues (1998: 43) as "information that is fit for use by information consumers". This definition clearly indicates the importance of the information relevancy for the data consumer. In this sense, there should not be any difference between the required and acquired information (Gerkes, 1997). To understand the level of information quality, a novel model was developed by Delone and Maclean (1992). Delone and Maclean (1992) categorized the four dimensions that affect the information quality as i) Intrinsic Information Quality, ii) Contextual Information Quality, iii) Representational Information Quality and iv) Accessibility Information Quality. According to Delone and Maclean (1992), the characteristics of the information quality models are illustrated (Figure 3.1).

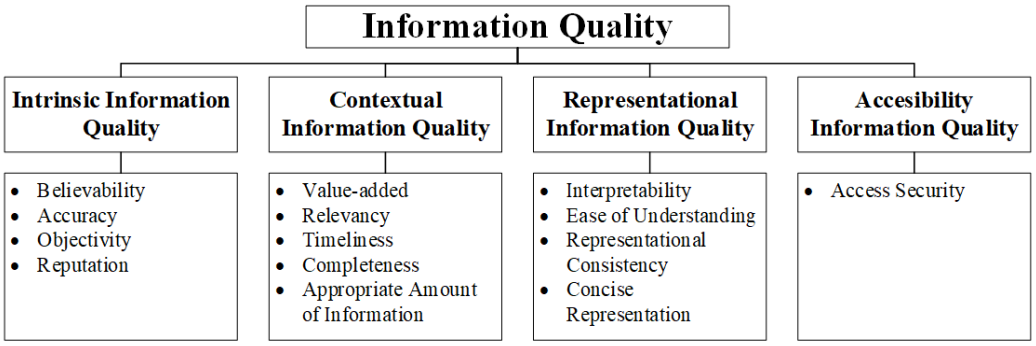


Figure 3.1. Conceptual Framework of Data and Quality (Wang and Strong, 1996).

As could be seen from Figure 3.1 that, representational information quality focus on interpretability, ease of understanding, representational consistency, and concise representation. Ease of understanding (also referred as understandability) is defined by Al-Hakim (2007:279) as "the degree to which the information can be comprehended by the user". In this context, the quality of safety information could be analyzed through ease of understanding. The conventional PowerPoint-based lectures contain presentation of visual information with verbal explanation (Bartsch and Cobern, 2003). PowerPoint-based lectures involve graphical elements during presentation, which supports the trainees to recall information (ChanLin, 1998; Szabo and Hastings, 2000). However, lectures are instructor-centered methods, which restrict the participation of trainees in the process of learning (Daluba, 2013). In this sense, the PowerPoint-based lectures could be suitable for trainees to learn necessary safety information, while, redundant information could lead to misunderstandings. To bridge this gap, the previous literature (Shakhashiri, 1984) suggests that lectures should involve necessary demonstrations. The demonstration is a two-staged procedure that the instructor shows how to do something in front of the students or to explain the principle as the first stage (Shakhashiri, 1984). As the second stage of demonstration, students practice the demonstrated content by doing on their own (Shakhashiri, 1984). In this sense, demonstrations enable trainees to consolidate their understanding through practice. To provide a compelling demonstration of safety content, V-SAFE provides an interactive environment to the trainees, which involves both individual and collaborative tasks. Therefore, they were able to improve their understanding through hands-on practice.

V-SAFE establishes a base for trainees to complete second stage of demonstration method, which enables conveying the safety information provided by PowerPoint-based lectures. The lectures contain not only basic safety information required but also the necessary demonstration of all activities, which take place in V-SAFE training. In the lights of this research, following hypothesis is developed: *Hypothesis 1.1: V-SAFE provides effective demonstration process for the trainees, thus, significantly improves the understandability of safety information, when used after PowerPoint presentations.*

Another critical parameter for the efficiency of safety information transfer process is implementing proper mnemonic strategies into the training process. Mnemonics are "any procedure or operation designed to improve ones' memory. (Scruggs *et al.*, 2010:1)". Therefore, to establish a base for the trainees to improve their learning level, as an initial step, memorizing the basics of safety information is a vital process. In a conventional safety training process, safety information is delivered through text-based notes, lectures or videos. All these techniques are based on visualizing the safety information through two-dimensional (2D) materials, which cause challenges in translating necessary safety information into a mental picture (Hadikusumo and Rowlinson, 2002). Developing visualization-based simulations that contain construction tasks and components could be a suitable solution for translating safety information into a mental image (Albert *et al.*, 2014). Thus, trainees could improve their mnemonics through spatial mnemonics strategy. Spatial mnemonics are methods that "uses the locations of fixed reference objects to learn information by associating and linking what you already know to what you wish to learn (Khan, 2016:161)". A seminal study by Krokos and colleagues (2018) found that virtual reality displays provide spatial awareness by leveraging the vestibular and proprioceptive senses. To test, whether virtual environments are suitable to improve spatial mnemonics just like virtual reality display sets, following hypothesis is developed: Hypothesis 1.2: V-SAFE significantly improves the spatial mnemonics of trainees when used after PowerPoint presentations. The prepared statement and hypotheses are illustrated in Figure 3.2.

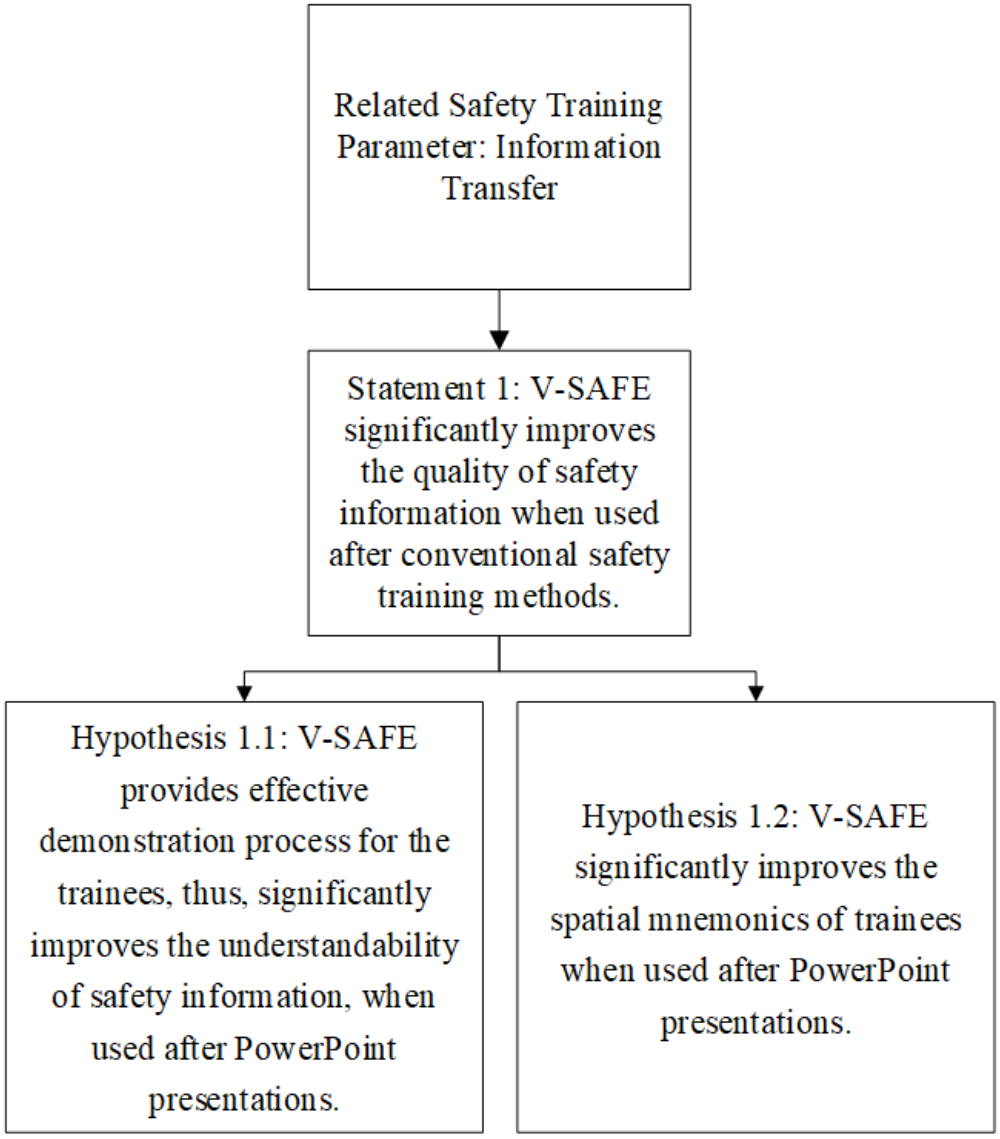


Figure 3.2. Hypothesis Related to Information Transfer.

3.1.2. Statement and Hypotheses Related to Instructional Feedback, Knowledge Development

An important parameter of effective safety training is to provide an effective feedback intervention to the trainees by instructional feedback mechanism (Burke *et al.*, 2006). An effective feedback method allows learners to correct their mistakes (Burke *et al.*, 2006). Feedback is also a vital characteristic of the programmed instructions, which are methods of training designed to represent the required information in a standardized manner, such as on a personal computer or in a workbook format (Burke

et al., 2006). In this sense, the contextual information quality could be considered as an essential parameter of effective feedback interventions. Contextual information quality refers to value-added, relevancy, timeliness, completeness, and the appropriate amount of information. Burke and colleagues (2006) state that programmed instructional feedback should be specific for each trainees' weaknesses or knowledge level. This statement could be considered as consistent with relevancy and the appropriate amount of information parameters defined by Wang and Strong (1996), since, the amount of information should be specific for each trainees' weaknesses. As a result of the feedback intervention, learners change their responses such as (i) reactions, (ii) learning of skills and knowledge and (iii) changes in the behavior (Kirkpatrick, 1979). Hence, the adequate value could be added to the trainees by relevancy in feedback, which are essential parameters of contextual information quality (see Fig. 3.1). In the lights of this context, by providing programmed instructional feedback effectively, trainees should be able to learn from their mistakes. Moreover, if they learn from their mistakes, then they improve their safety knowledge level by repeatable practice experience. According to the previous literature, adequate feedback (Kirkpatrick, 1979; Burke *et al.*, 2006) and effective safety training methods lead to the changes in the learners' behavior (Vinodkumar and Bhasi, 2010). In the lights of this research, the following hypothesis is developed: Hypothesis 2.1: *Trainees using V-SAFE significantly correct their mistakes by repeatable practice.*

Another critical parameter for effective safety training is knowledge development (Burke *et al.*, 2006). A considerable amount of literature (e.g., Nonaka, 1994; Egbu and Robinson, 2005; Nazim and Mukherjee, 2016; Nonaka and Nishiguchi, 2000; Carrillo and Chinowsky, 2006; Ellis, 2009) has been published on types of knowledge. Nonaka, who researched the dynamic theory of creating knowledge, suggested two dimensions of knowledge development (1994). One of the aspects was derived from the difference between two kinds of knowledge; tacit knowledge and explicit knowledge (Egbu and Robinson, 2005). The term explicit knowledge is defined as "formal and systematic knowledge, which can be expressed in words or numbers and can be documented or stored in databases as electronic records (Nazim and Mukherjee, 2016:50)". The explicit knowledge type differentiates from tacit knowledge about communication ease,

since, this kind of knowledge can be communicated between individuals formally and systematically. In safety training context, all information-based techniques such as pamphlets, lectures, safety notes are some methods for communicating explicit knowledge. As another significant knowledge type, tacit knowledge, which focuses on the experience of people, communicated in human-based activities as assessment, dispositions, perspectives, commitment, and motivation (Nonaka and Nishiguchi, 2000). Consequently, tacit knowledge has a personal quality which makes it difficult to formalize and communicate. Carrillo and Chinowsky also emphasized that tacit knowledge is stored in peoples' heads and acquired through experience; more difficult to document (2006). In safety training context, the general safety knowledge of a safety manager, the experience of a crane operator in commanding cranes are some examples of tacit knowledge. As another critical knowledge type, implicit knowledge is a bridge between tacit and explicit. The previous literature (Ellis, 2009) highlights that implicit knowledge is developed through the practical applications of explicit knowledge. Improving the implicit learning through repetitive practice could constitute tacit knowledge. For instance, an inexperienced safety engineer learns the essentials of the safety on the construction site, primarily in lecture-based presentation, safety notes or pamphlets. Therefore, he/she could develop adequate explicit knowledge. Later on, by monitoring the activities and exploring unsafe acts on the real construction site and practice the explicit knowledge acquired by conventional methods, he/she could develop necessary implicit understanding. Besides, if the safety engineer continuingly enters to the construction site and investigate unsafe situations repetitively, he/she gains adequate tacit knowledge through experience. In the safety training context, information-based techniques such as lectures, safety notes could be able to communicate necessary safety information, thus, developing explicit knowledge. However, these methods do not involve any practice; therefore, trainees could not be able to create essential implicit knowledge. To establish a base for the trainees to establish necessary implicit knowledge, V-SAFE environment provides simulation-based active practice. To check, whether trainees could develop implicit knowledge through V-SAFE, following hypothesis is developed: *Hypothesis 2.2: V-SAFE significantly improves implicit knowledge level of trainees by providing repeatable hands-on practice.* The developed statement and hypotheses are illustrated in Figure 3.3.

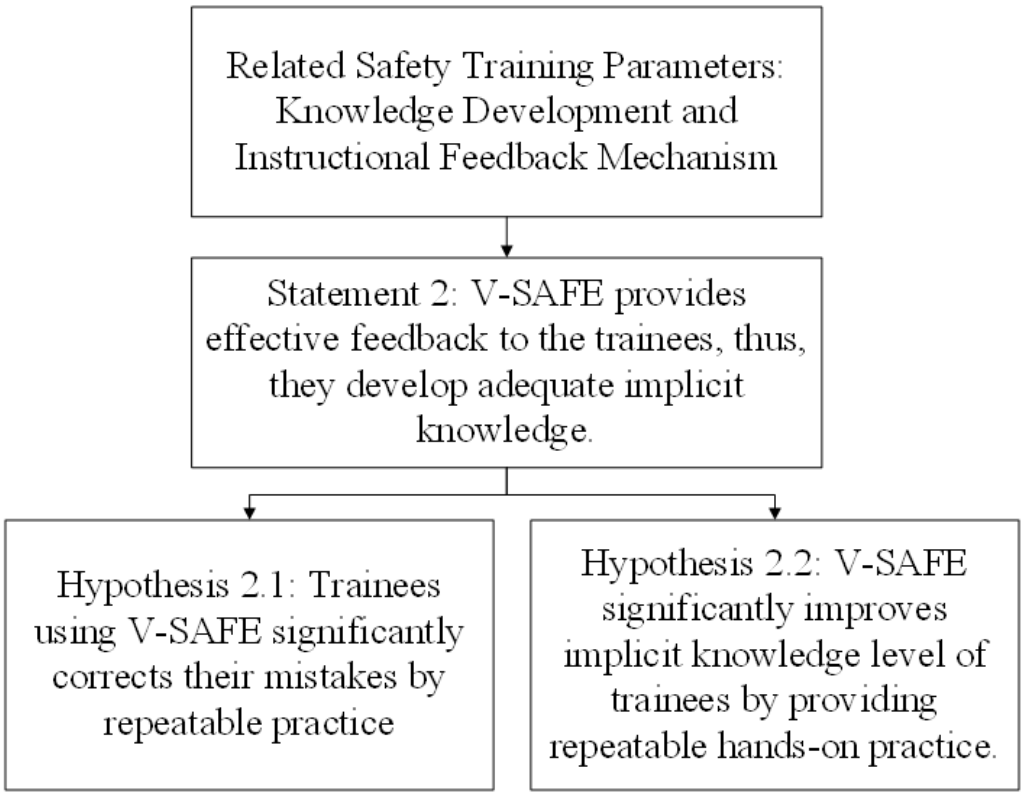


Figure 3.3. Hypotheses Related to Knowledge Development and Instructional Feedback.

3.1.3. Statement and Hypotheses Related to Hazard Identification

The previous literature (Patrucco *et al.*, 2010; Lingard and Rowlinson, 2005) argues that one of the main reasons for the on-site accidents is low hazard identification level. For this reason, an effective construction safety training method should establish a base for the trainees to identify on-site hazards effectively. The effectiveness of the hazard identification process should be evaluated by defining the risky situations that lead to fatal accidents. In this sense, assessing the trainees' hazard identification level through previously occurred accidents could be an appropriate approach. In order to measure the effectiveness of V-SAFE in each accident types, the following statement and hypotheses are developed (Fig. 3.4).

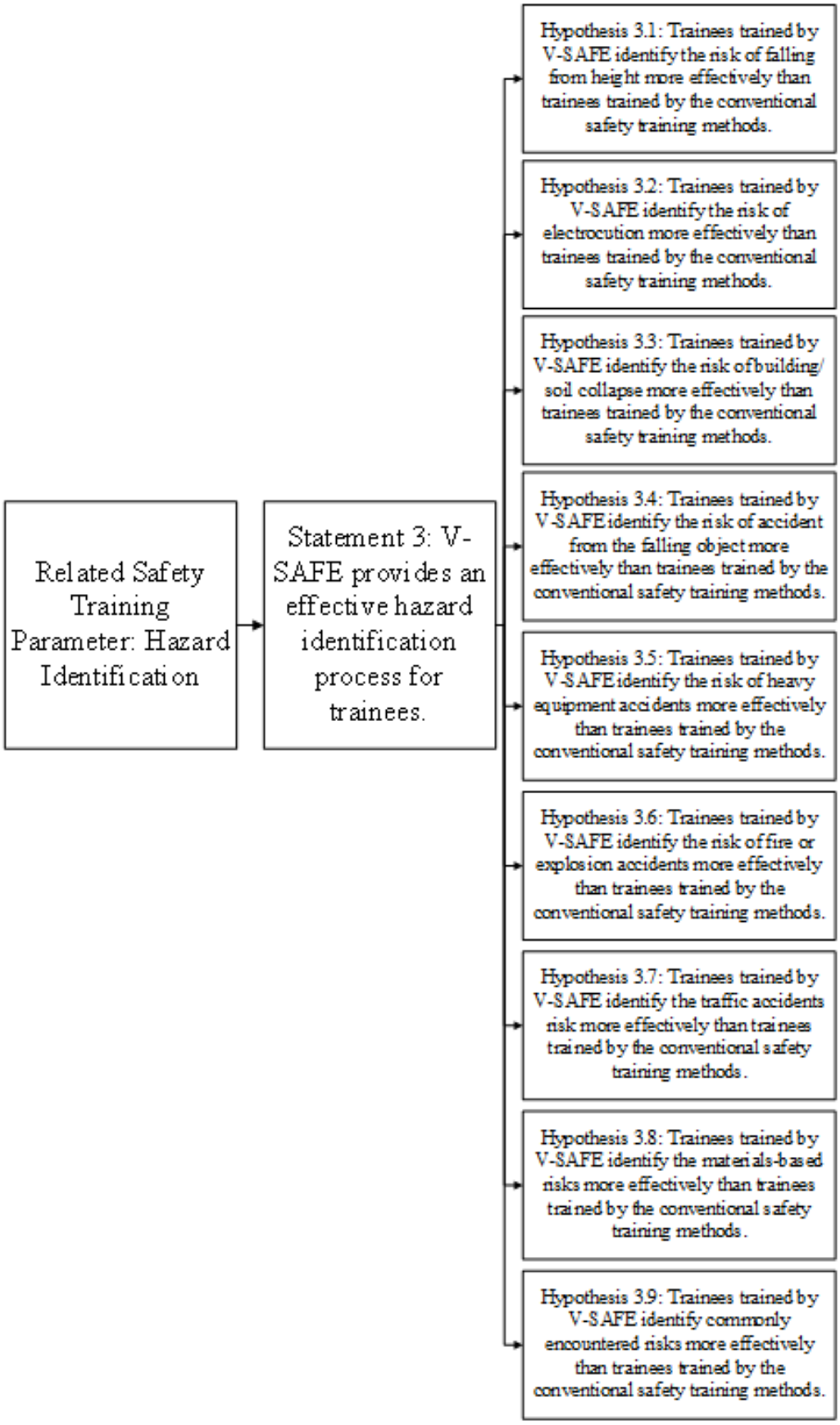


Figure 3.4. Hypotheses Related to Hazard Identification.

3.1.4. Statement and Hypotheses Related to Hands-On Practice

In a semantic study, Flick (1993:9) defines the term hands-on practice as "In terms of a specific instructional strategy where trainees are actively engaged in manipulating materials". Flicks' definition shows that hands-on training is crucial in experiential learning. In high-risk cases, providing hands-on practice could be dangerous, and a potential error of a trainee could result in a hazardous situation (Sisson, 2001). The most crucial parameter for effective hands-on practice, providing a similar on-site training process is quite essential. In other words, the practice should be completed under realistic working conditions and, as an ideal case, the trainee should do precisely the same practice, what they will be required to do every day (Sisson, 2001). In this sense, using simulations is quite suitable in the risky processes; thus, the trainee pretends to do the daily activities (Sisson, 2001). Therefore, when the instructor finds the performance of trainees' satisfying, then, they move on to the real job (Sisson, 2001). Klahr and colleagues (2007) argue that virtual technologies have a high potential to provide risk-free hands-on training and avoid the disadvantages of physical hands-on training such as risk factors. Previous literature (Sisson, 2001) states that a practical hands-on training method provides a realistic risk-free training environment to the real construction site. To evaluate the hands-on training level, comparing the safety behavior of the trainees on the actual construction site and virtual environment could be considered as a suitable method. In general, on-site hazard recognition is accomplished by monitoring field activities, thus, identifying risky situations during these activities (Albert *et al.*, 2014). Therefore, one can conclude that if the eye-tracking behavior of the trainees during virtual environment interaction is more effective in terms of eye-tracking metrics, V-SAFE provides an effective hands-on practice to the trainees.

In his seminal study, Wang and colleagues defined the factors that may influence safety risk tolerance (2016). Wang and colleagues state that one of the most critical elements to provide on-site safety is to sensitivity to the potential risks, which focuses on the capability of the workers to make quick response and judgment to potential threats. Thus, workers could be more sensitive to potential threats by taking immediate and correct safety action. For instance, when a worker detects the possible collapse of a

retaining wall, then, he or she will leave the area very quickly. As a result, a potentially fatal accident could be avoided. From this simple example, one could be concluded that the duration to take decisions is very significant to prevent a potential crash accident. In the eye-tracking context, time to first fixation refers to the "amount of time that it takes a respondent to look at a specific AOI from stimulus onset" (Farnsworth, 2018:5). In the light of this research, lower time to first fixation could be considered as a more effective safety behavior in construction safety management context, since, the lower amount of time to look at a trigger visual provides trainees to take quicker action. Therefore, following hypothesis is developed: *Hypothesis 4.1: The time to first fixation durations of participants using V-SAFE are significantly lower than the time to first fixation durations of the same participants on the real construction site.*

Another critical indicator to provide on-site safety is the workers' level of attention. Previous literature (Garrett and Teizer 2009; Rozenfeld *et al.* 2010) states that one of the main human-related factors that lead to on-site accidents is the lack of attention of workers when detecting potential hazards. Consequently, the workers could not be able to react correctly and take the appropriate decision (Garrett and Teizer 2009; Rozenfeld *et al.* 2010). In the eye-tracking context, total fixation duration (also known as time spent) "often indexes motivation and top-down attention, since respondents have to blend out other stimuli in the visual periphery that could be equally interesting". In this sense, due to the higher total fixation duration points out greater attention of the trainees, higher total fixation duration is a more preferred situation in the construction safety training context. Therefore, the following hypothesis is developed: Hypothesis 4.2: *The total fixation durations of participants using V-SAFE are significantly higher than the total fixation durations of the same participants on the real construction site.*

In their comparative analysis, Habibnezhad and colleagues (2016) evaluated the risk perception of workers impact to their visual search strategies when identifying hazards. The analysis results indicate that the trainees' with higher risk perception have higher first-fixation duration. In this sense, the following hypothesis is developed: *Hypothesis 4.3: The first fixation durations of participants using V-SAFE are significantly*

higher than the first fixation durations of the same participants on the real construction site.

Additionally, Hasanzadeh and colleagues (2017) analyzed the hazard identification ability on their attentional distributions by using eye-tracking metrics. Study results show that higher numbers of fixation count had the most significant impact on the hazard identification performance of the trainees. In this sense, the following hypothesis is developed: Hypothesis 4.4: *The fixation counts of participants using V-SAFE are significantly higher than the fixation count of the same participants on the real construction site.* The illustration of statement and hypotheses related to hands-on practice is represented in Figure 3.5.

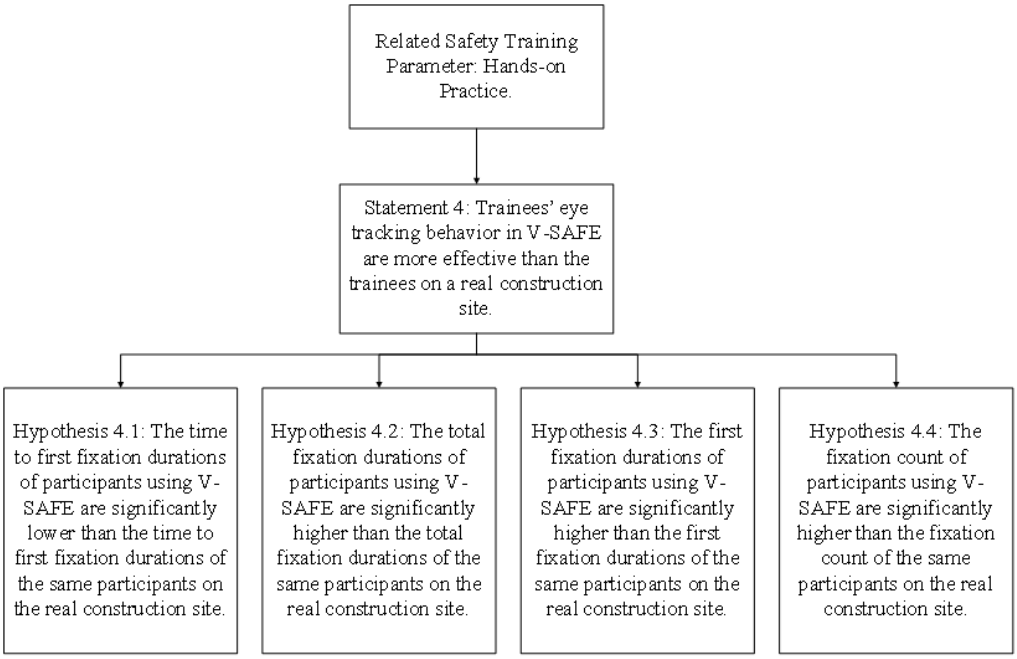


Figure 3.5. Hypotheses Related to Hands-On Practice.

3.1.5. Additional Related Statement and Hypotheses

Construction workers usually learn the essentials about safety and gather adequate hands-on practice during the tasks on the real construction site, since, conventional safety training methods fail to provide hands-on practice (Burke *et al.*, 2006). In this sense, the workers who do not pass a certain amount of time on the real construc-

tion tasks could be more likely to be subjected to on-site accidents. Consistently, Chi and colleagues (2005) highlight that on-site personnel with less than 1-year experience are exposed to around 80% of all fatal accidents, while the frequency of accidents is inversely correlated to the on-site work experience. To fill in this gap, previous studies (e.g. Guo *et al.*, 2012; Li *et al.*, 2012a; Li *et al.*, 2012b; Lin *et al.*, 2011; Mohd *et al.*, 2019; Kayhani *et al.*, 2019) have shown that virtual construction safety training methods have a high potential to provide appropriate hands-on training. However, these studies do not focus on the impact of site experience to the training effectiveness. In other words, there have been no controlled studies which compare the efficacy of virtual safety training based on trainees' site experience. To check, whether site experience of trainees affects the safety training performance, following hypothesis is developed: *Hypothesis 5.1: Both trainees with and without site experience could use virtual construction safety training tool in a similar and effective manner.*

Secondly, previous literature (e.g., De Kort and Ijsselsteijn, 2008; Nacke *et al.*, 2010) highlight that the effectiveness of virtual training methods is directly related to the computer gaming experience of trainees. In other words, previous studies (e.g., De Kort and Ijsselsteijn, 2008; Nacke *et al.*, 2010) argue that the trainees with higher computer gaming experience perform better compared to the trainees with less gaming experience. Indeed, construction workers usually have very limited computer literacy, including digital gaming experience. As workers with limited computer literacy are the target group of V-SAFE, it is essential to evaluate the impact of gaming experience on the effectiveness of V-SAFE training. In the lights of the research, following hypothesis is developed: *Hypothesis 5.2: Both trainees with and without computer gaming experience could use virtual construction safety training tool in a similar and effective manner.* The developed statement and hypotheses were illustrated in Figure 3.6.

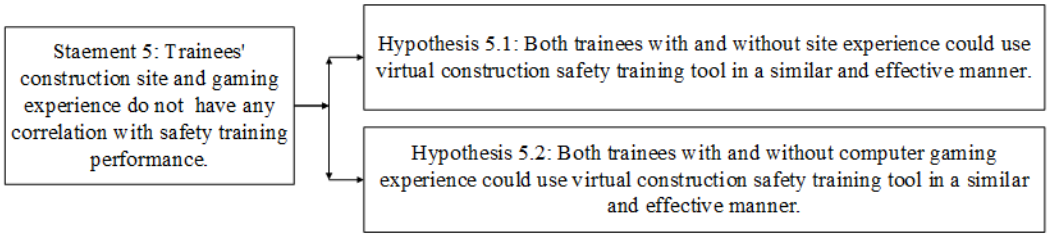


Figure 3.6. Additional Hypotheses Related to Effective Safety Training.

3.2. Tools and Parameters to Use

3.2.1. Virtual Environment: V-SAFE (Virtual Safety Analysis For Engineering applications)

One of the objectives of using the virtual technologies-based simulation is to model a real-time event or a hypothesis via a virtual environment to evaluate the deficiencies in a real system and to understand how the system works (Banks *et al.*, 2000). The use of the virtual simulation is suitable when interacting with the real system is not possible because of the non-accessibility to the high-risk system, or on developing a system (Sokolowski and Banks, 2011). In this sense, virtual environments provide a great opportunity for off-site training by enabling them to learn from their mistakes and to correct them without entering to the actual construction site. As a result, trainees could improve their behavior-based skills, communication, and cognitive abilities (Sherman and Craig, 2003). The simulation-based computer technologies have significant potential to improve the training level and to substitute conventional construction safety training methods.

In V-SAFE, a simulation-based virtual environment is developed to provide specific safety training for the tower crane tasks. In the very first step, V-SAFE simulates a tower crane lifting operation in a 3D virtual environment. V-SAFE is developed using the Unreal Engine 4. Unreal Engine is an open-source game engine developed by Epic Games for high-end games on personal computers (PCs), game consoles, and virtual reality (Unreal Engine Official Website, 2011). The Unreal Engine is not only used for PC gaming, but also used for other purposes such as education, training, transportation, movie storyboard, and simulation. Warner Bros, Sony, Amazon are some of the world-known licensees of the product (Unreal Engine Official Website, 2011).

The model framework of the V-SAFE involves five major stages. In the first phase, the tool generates the 3D models necessary for the simulation, such as buildings, bricks, personal protective equipment. In addition to that, the models in the surrounding environment, such as clouds, trucks, walls, are also developed. Thus, users could

monitor a realistic construction workspace in a 3D environment. Secondly, users enter the environment by creating their avatars on the server (Figure 3.7) and select one of the roles in the crane operation simulation; (i) crane operator, (ii) pointer and (iii) bricklayer. In the third stage, when all users select their role, they start the training by entering the warehouse, and they choose the correct personal protective equipment (PPE) (Figure 3.8). As the fourth step, the trainees start interacting with the models (e.g. crane, lift, rope, etc.) and other trainees in the virtual environment. Based on their behavior, decisions, and collaboration, their performances are analyzed, and they receive written feedback at the last stage of the training session.



Figure 3.7. An Avatar with PPE.



Figure 3.8. PPE Selection.

The main procedure of the training consists of three stages. Firstly, the potential hazards of tower crane tasks defined by Shepherd and colleagues are evaluated (2000). Moreover, major accident precursors during the crane operations such as blind lifts, load types, the wind, weather conditions, etc. are integrated into the virtual media. Later on, a feedback criterion is proposed that evaluates i) the safety knowledge and behavior of the trainees, ii) collaboration skills and iii) the level of conveying the safety

information. Subsequently, hazards are embedded into the virtual objects existed in the V-SAFE. As a result of literature review (e.g., Goh and Ubeynarayana, 2017; HSE, 2018; SSI, 2019b), the risks encountered during V-SAFE simulation were determined, and the potential hazards defined in the literature review are integrated into the V-SAFE via visualizations (Figure 3.9).



Figure 3.9. Visual risk sources integrated into V-SAFE.

As the vital principle of the V-SAFE training process, a scenario was developed that focuses on laying up a brick wall on the second floor of a building, without causing any accidents. In the scenario, the materials should be safely loaded to the crane and transported to the second floor, which is a common practice on the real construction sites. To accomplish this task, all trainees should fulfill both collaborative and personal responsibilities, which are different for each role. There are three different roles in the scenario as (i) crane operator, (ii) pointer and (iii) bricklayer. Crane operators' primary duty is to move the load from ground to the second floor of the building, without causing any safety issues. Pointers' primary responsibility is to guide the crane operator during the transportation of the load, especially in the blind spots.

The bricklayer is responsible for loading and unloading processes. During the V-SAFE training, potential accidents might occur based on the safety behavior of trainees. In other words, the inappropriate safety behavior of trainees may lead to accidents in the virtual environment. To establish a base for the trainees to learn from their mistakes, V-SAFE provides specific feedback for each trainee based on their behavior after the training process. The training framework of V-SAFE is illustrated in Figure 3.10.

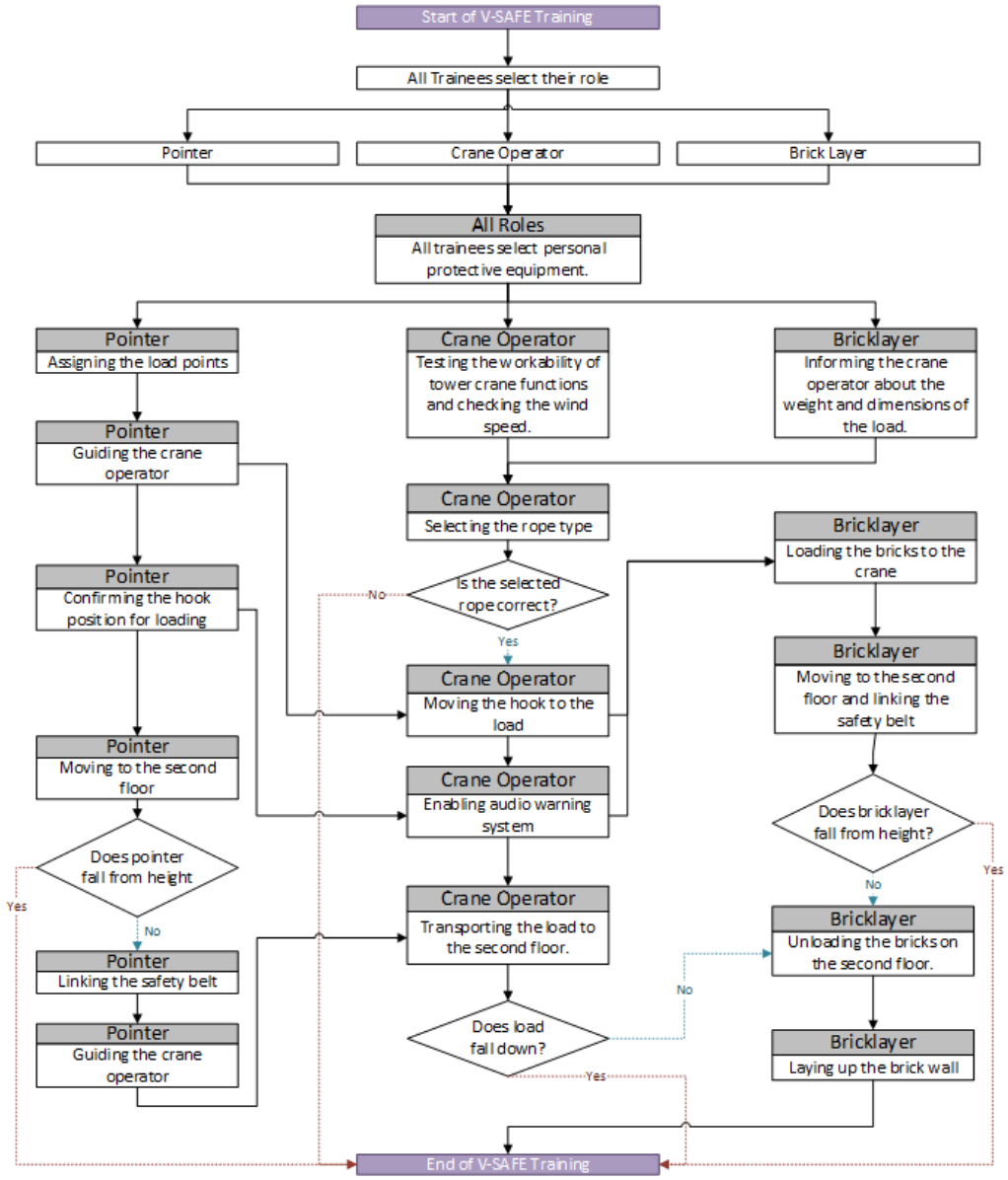


Figure 3.10. V-SAFE Training Framework.

The system architecture of V-SAFE is based on the modularization of three layers as (i) resource, (ii) content and (iii) simulation. The system architecture of V-SAFE

and relationships among the three layers are illustrated in Figure 3.10. As could be seen from Fig. 3.11 the information flow in V-SAFE is sequential, and a unilateral relationship occurs between the layers.

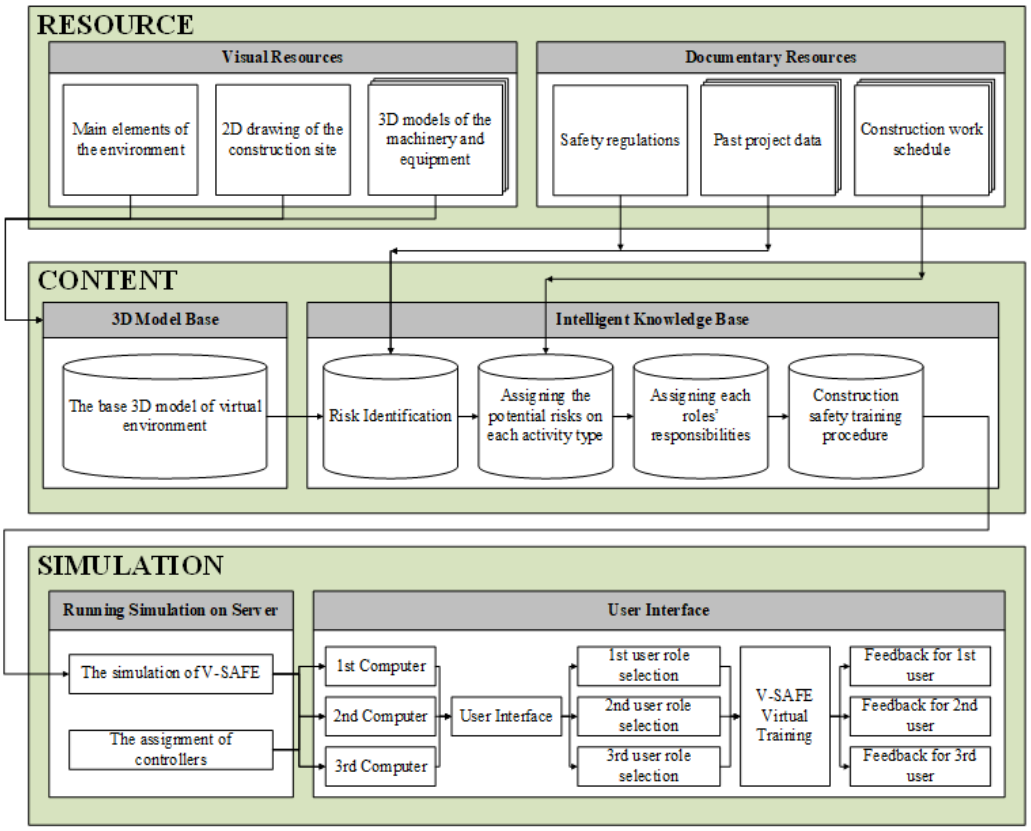


Figure 3.11. System Architecture of V-SAFE.

The resource layer includes two sub-resources as visual and documentary resources. The visual resources establish a base for the 3D model in the content layer, while, the documentary resource layer consists of construction work schedule involves tower crane tasks, past project data, and safety regulations. The primary usage of the documentary resources is to provide necessary data for both the 3D model and intelligent knowledge bases. In brief, the main function of the resource layer is to store the necessary data on the cloud server and to transfer the data to the relevant elements of the content layer.

Secondly, the content layer is another crucial part of V-SAFE architecture. There are two resources in the content layer including 3D model base and intelligent knowledge

base. The major function of the 3D model base is to visualize the virtual environment of V-SAFE and to locate all equipment and machinery used during the tower crane tasks. In general, the risk identification process is conducted through monitoring hazardous materials and risky situations in the virtual space. For instance, if users monitor an electric cable on the virtual construction site, they would easily grasp potential electrocution risk. Therefore, the 3D model base is a valuable information resource of the risk identification process enlisted in the intelligent knowledge base. The intelligent knowledge base is designed to transform the gathered data from the documentary resources and the 3D model base, later on, compile this processed data to the simulation module. In other words, an intelligent knowledge base creates knowledge through analyzing the data transferred from other modules and conveys the established knowledge to the necessary module. Due to the selective transmission of the data, the knowledge base could be considered as "intelligent". In addition, information about potential accidents is stored in an intelligent knowledge base. According to the simulation settings, if an accident occurs during the virtual training, simulation will be automatically terminated. To summarize, the main function of the content layer is to establish a bridge between the resource and simulation modules.

Thirdly, simulation is the final module of V-SAFE architecture. There are two sub-parameters in the simulation module, which are running simulation on the server and user interface. The main objective of using the servers is to provide a multiplayer platform and to organize the simulation of V-SAFE smoothly. Later on, users could select their role to start the simulation. Therefore, all users could enter to the simulation of V-SAFE. After the completion of the V-SAFE training session, each user receives specific feedback based on their performance.

3.2.1.1. Input Devices and Gaming Controllers. Many virtual environment-based applications such as PC games, serious games use standard controllers such as joysticks, joypads or touchpads. The significant advantages of these controllers are low-cost, easy configuration, ease to implement, and use. Besides, many trainees with gaming experience are quite familiar with these controllers in their daily lives. Therefore, it is aimed

to increase the physical interaction quality between the trainees and the virtual environment by improving the usability of the simulation system. Trainees could interact with the virtual environment by using joysticks (Logitech Extreme 3D Pro Joystick), gamepads (Logitech Gamepad F310), and standard interaction method (keyboard and mouse).

During the locomotion of the real tower cranes, the movement of the hook is commanded through the joystick orientation (Figure 3.12). To provide an enhanced simulation experience during the crane operations, air flight simulator joysticks are used, since, it is determined that, air flight simulator joysticks are quite similar to tower crane operators' controllers (Figure 3.13).



Figure 3.12. The interface of the crane operator in a real tower crane.

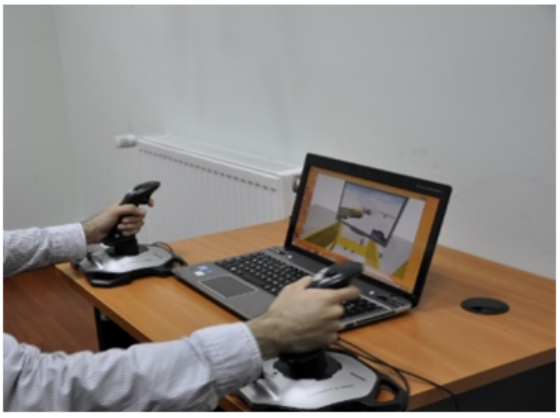


Figure 3.13. The controller setting of trainees during V-SAFE training.

Logitech Extreme 3D Pro Joystick is a joystick controller, which is developed for air flight simulation. Rudder Control, Programmable twelve buttons, 8-way hat switch are some of the specifications of Logitech Extreme 3D Pro Joystick (Logitech Official Website, 2019a). In a study occurred in the construction safety management literature (e.g., Rogers *et al.*, 2017) used Logitech Extreme 3D Pro Joystick for the locomotion of the tower cranes.

Logitech F310 is a twelve button gamepad developed by Logitech. It contains 4-switch standard D-pads rest on a single pivot point, resulting in mushy control (Logitech Official Website, 2019b). The D-pad glides over four individual switches for a more responsive, tactile feel (Figure 3.14). During the conducted experiments, some trainees found it difficult to use the keyboard and mouse layout, so they selected gamepads, which provide them a simpler interface (Figure 3.15).



Figure 3.14. Logitech F310.



Figure 3.15. The interaction of the trainees with V-SAFE via Logitech F310.

3.2.2. Eye Tracking Devices

Several eye metrics are defined in eye-tracking terminology. Firstly, the Area of Interest (AOI) is an important parameter to identify the eye-tracking behavior of the trainees. AOI defined by Farnsworth (2018:3) as "A tool to select subregions of the displayed stimuli, and to extract metrics specifically for these regions". Therefore, AoI could be assigned as considering the objects, points, or regions that are crucial. For instance, a stair in working at height activity could be regarded as vital to identify fall accidents. In this sense, stairs could be assigned as AOI. Secondly, gaze (also referred as dwell, fixation cluster) "is usually the sum of all fixation durations within a prescribed area. It is best used to compare attention distributed between targets. It can also be used as a measure of anticipation in situation awareness if longer gazes fall on an area of interest before a possible event occurring" (Mello-Thoms, 2002:113). Major metrics are also defined in Table 3.1.

Table 3.1. The summary of eye movement metrics.

Term	**Definition**	**Unit**
General Eye-Tracking Terms		
Fixations and Gaze Points	"Gaze points show towards which elements of a stimulus the eyes are looking at. If a series of gaze points is very close ? in time and in space - this gaze cluster constitutes a fixation, denoting a period where the eyes are locked towards an object" (Ballco and de-Magistris, 2017:5).	-
Saccade	"The eye movements between fixations are generally referred to as saccades". (Ballco and de-Magistris, 2017:5).	-
Areas of Interest (AOI)	"A tool to select subregions of the displayed stimuli, and to extract metrics specifically for these regions" (Farnsworth, 2018:3).	-
Eye-Tracking Metrics		
Total fixation duration (also referred to as time spent or dwell time)	"Time spent" quantifies the amount of time that respondents have spent on an AOI. Time spent often indexes motivation and top-down attention since respondents have to blend out other stimuli in the visual periphery that could be equally interesting. Long prevalence at a certain region clearly indicates a high level of interest, while shorter prevalence times indicate that other areas on screen or in the environment might be more interesting" (Farnsworth, 2018:6).	Second
First fixation duration	"The duration of the first contact with the AoI when more than one fixation is made on the related AOI" (Dolgunsöz, 2015:89)	Second
Time to first fixation	"The time to first fixation indicates the amount of time that it takes a respondent (or all respondents on average) to look at a specific AOI from stimulus onset" (Farnsworth, 2018:5).	Second
Total Visit Duration	"Total Visit Duration is defined as the sum of visit durations of an active AOI" (or AOI group). An individual visit is defined as the time interval between the first fixation on the active AOI and the end of the last fixation within the same active AOI where there have been no fixations outside the AOI" (Tobii Pro, 2014:105).	Second
Visit Count	"A visit is defined as the time interval between the first fixation on the active AOI and the end of the last fixation within the same active AOI where there have been no fixations outside the AOI" (Tobii Pro, 2014:107).	#
Fixation Count (also referred as number of fixations)	"Number of eye fixations detected during an interval of interest" (Toker *et al.*, 2013:4).	#

To gather the eye-tracking data of trainees, a Tobii X2-30 compact on-screen eye-tracker was used (Figure 3.16). Tobii X2-30 is a screen-based eye tracker capturing gaze data at 30 hertz. The technical specification of the Tobii X2-30 eye-tracker involves: accuracy of 0.4 degrees, precision of 0.32 degrees, freedom of head movement 50 centimeter (width) x 36 centimeter (height) x 90 centimeter (depth), cm (20 x 14”), system latency of 50 to 70 milliseconds range, 30 hertz of data rate and 9 points calibration (Tobii Pro, 2014). Secondly, Tobii Pro Glasses 2 wearable eye tracker was used during the on-site experiment (Figure 3.17). Tobii Pro Glasses 2 device was designed to capture natural viewing behavior in any real-world environment while ensuring outstanding eye-tracking robustness and accuracy (Tobii Pro, 2018). The technical specifications of Tobii Pro Glasses 2 are gaze sampling frequency of 100 hertz, 1 point calibration, scene camera recording angle of 82 degrees (horizontal) and 52 degrees (vertical).



Figure 3.16. Tobii X2-30.



Figure 3.17. Tobii Pro Glasses 2.

3.3. Data Analysis Methods

Accurate interpretation of the results of research depends primarily on the selection of the appropriate statistical test for the purpose and data type. The analyses in this study were conducted for various purposes. Besides, the types of variables measured and experimental setting differ from experiment to experiment. One group paired data, two groups of independent samples of factorial experiments are some examples of the different experimental settings. For this reason, the systematic selection of the analysis method used in the experiments has gained importance.

3.3.1. Comparing One Group for Two Related Samples

3.3.1.1. Paired Samples T-test. Paired samples t-test is a statistical analysis method to check if two paired data are identical (Maxwell *et al.*, 2017). Paired samples t-test is a suitable method if the same parameter is quantified under different conditions (Maxwell *et al.*, 2017). In general empirical analysis, one sample exists, which is usually referred to as the control data. Later on, statistical treatment is done, and the same measurements are retaken. In that case, the second data set referred to as the treatment data. The paired samples t-test mainly concentrates on the difference between the results of the control and treatment data. It is planned to use paired samples t-test, especially in the research that will be conducted with the same population. The main assumptions of paired samples t-test are described as the following:

- The Number of Groups: There must be two groups (Field, 2009).
- Normality: The distribution of differences between scores must be normally distributed (Field, 2009).
- Data Type: The measured data must be at the interval or ratio level (Field, 2009).
- Dependent Groups: The measured scores must be related to each other (Field, 2009).
- Homogeneity of Variances: The variances in the measured populations must be roughly equal (Lehman *et al.*, 2013).

- Outliers: The measures scores must not involve any significant outliers (Field, 2009).

3.3.1.2. Wilcoxon Signed-Rank Test. The comparison of the mean of more than two main populations was made by parametric paired samples t-test. In cases, where underlying assumptions such as normality and homogeneity are not provided for parametric tests, the non-parametric tests are alternative to these tests. The non-parametric form of paired sample t-test could be conducted through the Wilcoxon Signed-Rank Test. This test is used to analyze whether there are differences between observations of the same individuals (Büyüköztürk, 2004). The main assumptions of the Wilcoxon Signed-Rank Test are described as the following:

- The Number of Groups: There must be two groups (Taeger and Kuhnt, 2014).
- The Dependency of Groups: The measured scores must be related to each other (Taeger and Kuhnt, 2014).
- Data Type: The measured data must de at the interval, ratio level, or ordinal scale (Taeger and Kuhnt, 2014).
- Symmetrical Data: The random variables follow continuous distributions, which might differ, but are all symmetric about the same median (Taeger and Kuhnt, 2014).

3.3.1.3. McNemar Test. The McNemar test is used for determining whether there is a difference in the bilateral dependent variable between the two related groups (Riffenburgh, 1999). The method is quite similar to the paired t-test. However, the major difference of McNemar Test is the data type used in the method. McNemar Test uses nominal data. As a difference compared to paired samples t-test is that the data is discrete. The McNemar test is a widely used method of analyzing paired samples and case-control studies, as well as analyzing the latest pre-test study designs (Riffenburgh, 1999). At the point when a cell in a 2x2 possibility table has an estimation of 0 or 1, Wolf-Haldane changes were used (Lavado' Valenzuela *et al.*, 2011). The main assumptions of the McNemar Test are described as follows.

- The Dependency of Groups: The measured scores must be related to each other (Riffenburgh, 1999).
- Data Type: The measured data must de at categorical or nominal data (Riffenburgh, 1999).
- Randomness: The sample group must be random from the population (Riffenburgh, 1999).

3.3.2. Comparing One Group for Several Related Samples

3.3.2.1. One-Way ANOVA with Repeated Measures. The repeated measures is a research design involving multiple measurements of the same variable taken under different conditions on the same subjects (Salkind, 2010). One-Way ANOVA with repeated measures is a frequently used statistical approach in paired research designs (Gueorguieva and Krystal, 2004). As with all statistical analyzes, certain assumptions must be made to justify the use of this test. Violations can affect results and often cause inflation of the type 1 error (Green and Salkind, 2008). The main assumptions of One-Way ANOVA with Repeated Measures are described as the following:

- Data Type: The measured data must de at the interval, ratio level, or ordinal scale (Girden, 1992).
- The Number of Groups: There must be at least three groups (Girden, 1992).
- The Dependency of Groups: The measured scores must be related to each other (Girden, 1992).
- Multivariate Normality: Dependent variable must be approximately normally distributed (Girden, 1992, Vasey and Thayer, 1987).
- Sphericity: The differences in variances among the combinations of dependent groups must be equal (Girden, 1992).
- Outliers: In the related groups, outliers must not exist in the dataset (Girden, 1992; Vasey and Thayer, 1987).
- Randomness: The sample group must be random from the population (Girden, 1992).

3.3.2.2. Friedman Test. In the previous section, the One-Way ANOVA with repeated measures method is introduced. One-Way ANOVA with repeated measures has been found to have essential assumptions. If one of these assumptions is violated, it causes insignificance in the results. In these cases, there is another alternative for repeated measures: the Friedman Test. Therefore, in cases where there are more than two conditions, the use of the Friedman Test is suitable (St. Laurent and Turk, 2013). The main assumptions of the Friedman Test are described as the following:

- Data Type: The measured data must de at the interval, ratio level, or ordinal scale (St. Laurent and Turk, 2013).
- The Number of Groups: One group must be measured on three or more different conditions. (St. Laurent and Turk, 2013).
- The Dependency of Groups: The measured scores must be related to each other (St. Laurent and Turk, 2013).
- Randomness: The sample group must be random from the population (St. Laurent and Turk, 2013).

3.3.3. Comparing Two Groups for Independent Samples

3.3.3.1. Student T-test. The student T-test is a two population analysis method and generally used when; (i) sample size is small, (ii) datasets are normally distributed and (iii) two normal distributions are not known (Walpole *et al.*, 2007). The student t-test is a commonly used statistical method for comparing two independent sample spaces. The main assumptions of student t-test are described as the following:

- Normality: The distribution of samples must be normally distributed (Field, 2009).
- Interval Data: The measured data must be at the interval level (Field, 2009).
- Homogeneity of Variances: The variances in the measured populations must be roughly equal (Field, 2009).
- Mutually Exclusive Groups: The measures of the data must be independent of each other (Field, 2009).

- Randomness: The sample group must be random from the population (Field, 2009).

3.3.3.2. Mann-Whitney U test. Mann-Whitney U test is an alternative non-parametric form of student t-test. Mann Whitney U test is performed, if an independent variable with two subgroups and a dependent variable are of the ordered (continuous) data type (Nachar, 2008). The Mann-Whitney U Test is also an auxiliary test that can be used when the t-test cannot be performed for various reasons (Nachar, 2008). This test is, therefore, a non-parametric alternative to the t-test (Nachar, 2008). The main assumptions of Mann-Whitney U test are described as the following:

- Interval Data: The measured data must be at the interval level (Nachar, 2008).
- Mutually Exclusive Groups: The measures of the data must be independent of each other (Nachar, 2008).

3.3.4. Factorial Analysis

3.3.4.1. Two-Way ANOVA. One-Way ANOVA is a quite common statistical test method to compare the mean scores from different groups. Two-Way ANOVA is an expanded form of the One-Way ANOVA method, which evaluated the impact of two categorical variables on one dependent variable (Albright and Winston, 2014). Two-Way ANOVA provides a comparison between the mean differences of groups that are divided into two independent variables (factors) (Albright and Winston, 2014). The primary purpose of using Two-Way ANOVA is to determine the potential correlation between the dependent variable and two independent variables (Hesterberg *et al.*, 2005). In other words, using Two-Way ANOVA provides evaluating the impact of each independent variable on the dependent variable and the potential interaction between independent variables (Walpole *et al.*, 2007). The assumptions of Two-Way ANOVA are described as the following:

- Interval Data: The measured dependent variable must be at the interval level (Albright and Winston, 2014).
- Mutually Exclusive Groups and Categorical Variables: The measures of the data must be independent of each other and involve two or more categorical variables (Albright and Winston, 2014).
- Homogeneity of Variances: The variances in the measured populations must be roughly equal (Albright and Winston, 2014).
- Normality: The distribution of samples must be normally distributed (Albright and Winston, 2014).
- Randomness: The sample group must be random from the population (Albright and Winston, 2014).
- Outliers: In the related groups, outliers must not exist in the dataset (Albright and Winston, 2014).

3.3.4.2. Kruskal Wallis-H Test. Another ordinarily utilized method to evaluate one factor is Kruskal-Wallis H test (Weaver *et al.*, 2017). The Kruskal-Wallis H test is the non-parametric alternative simple to the parametric, One-Way ANOVA. The Kruskal-Wallis H test is additionally viewed as like the Mann-Whitney U test, as it plays out an examination or investigation of autonomous examples. The Kruskal-Wallis is most pertinent when looking at least two examples; the information inside every one of the different examples do not have to pursue a typical dispersion (Weaver *et al.*, 2017). The assumptions of Kruskal Wallis-H Test are described as the following:

- Interval Data: The measured dependent variable must be at the interval level (Weaver *et al.*, 2017).
- Mutually Exclusive Groups and Categorical Variables: The measures of the data must be independent of each other and involve two or more categorical variables (Weaver *et al.*, 2017).

3.3.4.3. Scheirer-Ray-Hare Test. The Scheirer-Ray-Hare test is a non-parametric form of Two-Way ANOVA test, which could be used for analyzing whether a measurement

is affected by two or more factors (Scheirer *et al.*, 1976). The method is an extension of the Kruskal-Wallis H test. As an addition to the Kruskal-Wallis H test, the method could analyze the impact of two or more factors and their interrelations (Scheirer *et al.*, 1976). Just like other non-parametric methods, the analysis method is based on evaluating the ranks of the samples, rather than the actual observations (Scheirer *et al.*, 1976). The assumptions of Scheirer-Ray-Hare Test are described as the following:

- Interval Data: The measured dependent variable must be at the interval level (Scheirer *et al.*, 1976).
- Mutually Exclusive Groups and Categorical Variables: The measures of the data must be independent of each other and involve two or more categorical variables (Scheirer *et al.*, 1976).

3.3.5. Tests to Evaluate Assumptions

3.3.5.1. Shapiro-Wilk and Kolmogorov-Smirnov Test for Normality. Shapiro-Wilk test is used to test, whether a randomly obtained sample data conforms to a certain distribution (uniform, normal or poison) (Walpole *et al.*, 2007). The null hypothesis of this test is that the population is normally distributed. Therefore, on the one hand, if the p-value is lower than the selected alpha level, the null hypothesis is rejected, and there is evidence that the tested data are not normally distributed. On the other hand, if the ρ-value is higher than the selected alpha level, the null hypothesis that the data comes from a population with a normal distribution cannot be rejected (Razali and Wah, 2011). To assess the normality, it is aimed to use the Shapiro-Wilk test, since, Shapiro Wilk test was originally restricted for the sample size of less than 50 (Shapiro and Wilk, 1965). In four of five conducted experiments, the sample size is less than 50. Therefore, rather than using alternative normality tests such as Kolmogorov-Smirnov Test, Anderson-Darling Test, it was determined to use Shapiro-Wilk Test in these four experiments, sample sizes lower than 50.

Secondly, the Kolmogorov-Smirnov is another normality test. In principle, the Kolmogorov-Smirnov test is based on the comparison of the sample data with the

proposed cumulative distribution function (Walpole *et al.*, 2007). With the support of this test, it is possible to examine whether the data collected from a sample exhibits normal distribution (Walpole *et al.*, 2007). In one of the conducted experiments, the sample size was significantly larger than 50. In this test, it is determined to use the Kolmogorov-Smirnov test.

3.3.5.2. Levene and Bartlett Tests for Homogeneity. Levene and Bartlett tests evaluate, whether the variances of the masses were equal. The Bartlett test assumes that the samples (groups) are independent and have a normal distribution of the sampled masses, and these are strong assumptions (Coşkun and Keskin, 2008). The assumptions of the Levene test are the independent distribution of the samples and the continuous distribution of the sampling populations, and the assumption of a continuous distribution is much weaker than the assumption of having the normal distribution of the Bartlett test (Coşkun and Keskin, 2008). The Bartlett test is included in the parametric tests class since the fact that the Bartlett test has a normal distribution assumption transforms the problem into the problem of equality of the variances of the masses with the normal distribution. However, the asymptotic distribution of the test statistic of the Levene test is not independent of the distribution for the family of continuous distributions (Coşkun and Keskin, 2008). The test of equality of variance is also referred as the homogeneity test of variance in the literature. Homogeneity tests are used as hypothesis controllers to use statistical tests (such as analysis of variance, t-test) that require the assumption of homogeneity in theory (Coşkun and Keskin, 2008). Since the distribution of two samples t-test statistics has different degrees of freedom t distribution according to whether the variance is equal or not under normal theory, the information about the homogeneity of variances in the two samples t-test plays an essential role in the choice of degree of freedom.

3.3.5.3. Wald-Wolfowitz Runs Test for Randomness. Wald-Wolfowitz is a nonparametric statistical test that has been tested (or simply tests) after statisticians Abraham Wald and Jacob Wolfowitz, checking the randomness hypothesis for a two-valued data set (Magel and Wibowo, 1997). Instead, it can be used to test the hypothesis that

the elements of the sequence are mutually independent (Magel and Wibowo, 1997). Wald-Wolfowitz runs test can be used to test the randomness of distribution by taking the data in the given order and marking the data larger than the median with plus, while, marking the data less than the median with a minus (Magel and Wibowo, 1997). Secondly, Wald-Wolfowitz runs test is suitable for evaluating, whether or not a function fits the data appropriately set by marking function value with plus and the other data with a minus (Magel and Wibowo, 1997).

3.3.5.4. Mauchly Test for Sphericity. Sphericity is an essential assumption of an ANOVA with repeated measures. This refers to the condition that the difference between all possible pairs of subject conditions are equal to each other. Violation of sphericity occurs when the differences between all combinations of conditions are not equal. If the sphericity is violated, then the variance calculations may be distorted, resulting in an inflated F ratio (Girden, 1992). Sphericity can be evaluated when there are three or more levels of a repeated measurement factor, and the risk of violating the sphericity increases with each repeated measurement factor. If sphericity is violated, it should be decided whether the univariate or multivariate analysis is selected. If a univariate method is selected, repeated measures should be appropriately corrected depending on the degree of ANOVA, the degree of sphericity (Greenhouse and Geisser, 1959). When sphericity is established, the F ratio is valid and can, therefore, be interpreted. However, if the Mauchly test is significant, the F-rates generated should be interpreted with caution as violations of this assumption may lead to an increased Type I error rate and affect the results from your analysis. Where Mauchlys' test is significant, changes in degrees of freedom are required to obtain a valid F ratio.

3.3.5.5. Multivariate Test of Normality. In the theory of probability and statistical sciences, the multivariate normal distribution or multivariate Gaussian distribution is the generalization of the normal distribution (or Gaussian distribution), which is a univariate distribution, to multiple variants (Cox and Small, 1978). If the X and Y random variables exhibit a normal distribution and are statistically independent of each other, then these two random variable compounds (i.e., the random vector) show

a bivariate normal distribution or, in other words, are normally distributed. However, it is not true that both random variables, which show a common normal distribution, are independent of each other (Cox and Small, 1978).

3.3.5.6. Cohen t-test for the Power and Effect Size Evaluation. Cohen (2019) defined statistical power as the probability of rejection of a false null hypothesis (the probability of type 2 error (β)). In Neyman-Pearson theory, the mathematical definition of statistical power is expressed as 1-β (Cohen, 2019). As can be seen from this statement, the probability of making a second type of error decreases, as statistical power increases. As a natural result of this; the statistical power value is between 0 and 1 (Sedlmeier and Gigerenzer, 1989:309). When determining the number of samples by power analysis, it is necessary to determine achieved power (Özdamar, 2003; Kirby *et al.*, 2002). Experimental posthoc power analysis is done at the end of the study, and the power of the analysis is evaluated to determine whether the results have a significant effect, based on the achieved power value. In practice, achieved power value with sensible defaults for some parameters could be assigned as a significance level of 0.05 and a power level of 0.80 (Ellis, 2010).

Effect size is the statistical value that shows the deviation from the expectations defined in the null hypothesis of the results obtained from the sample (Cohen, 2016; Vacha-Haasse and Thompson, 2004). Effect size is generally defined as the magnitude of the difference between the null hypothesis and alternative hypotheses. This is an indication of the practical significance of the research results. The effect size is the concept of how much different a new method makes compared to the old one and can be calculated in different ways (Kılıç, 2014). The effect size is a concept that must be taken into consideration when calculating the minimum required sample size when planning any study (Kılıç, 2014). It is possible to decide on the magnitude of the impact by reviewing the literature related to the subject and if there is not any study done or reached on this subject with the past experiences of the researcher (Kılıç, 2014). For a non-exemplary study, pilot-preliminary research can be conducted to determine the effect size (Kılıç, 2014). The larger the expected difference between the groups,

the larger the effect size, the smaller the sample size will suffice, since it will be easier to put forward the difference statistically (Kılıç, 2014). According to Cohen (2016), effect size impact range is defined as the following (Table 3.2). The method selection flowcharts are shown in Figure 3.17 and Figure 3.18.

Table 3.2. Levenes' test values and corresponding effect size (adopted from Cohen, 2013).

Levenes' test d value	Small Effect Size	Medium Effect Size	Large Effect Size
T-test	0.20	0.50	0.80
Mann-Whitney U Test	0.20	0.50	0.80
Kruskal-Wallis H Test	0.10	0.25	0.40
ANOVA	0.01	0.07	0.14

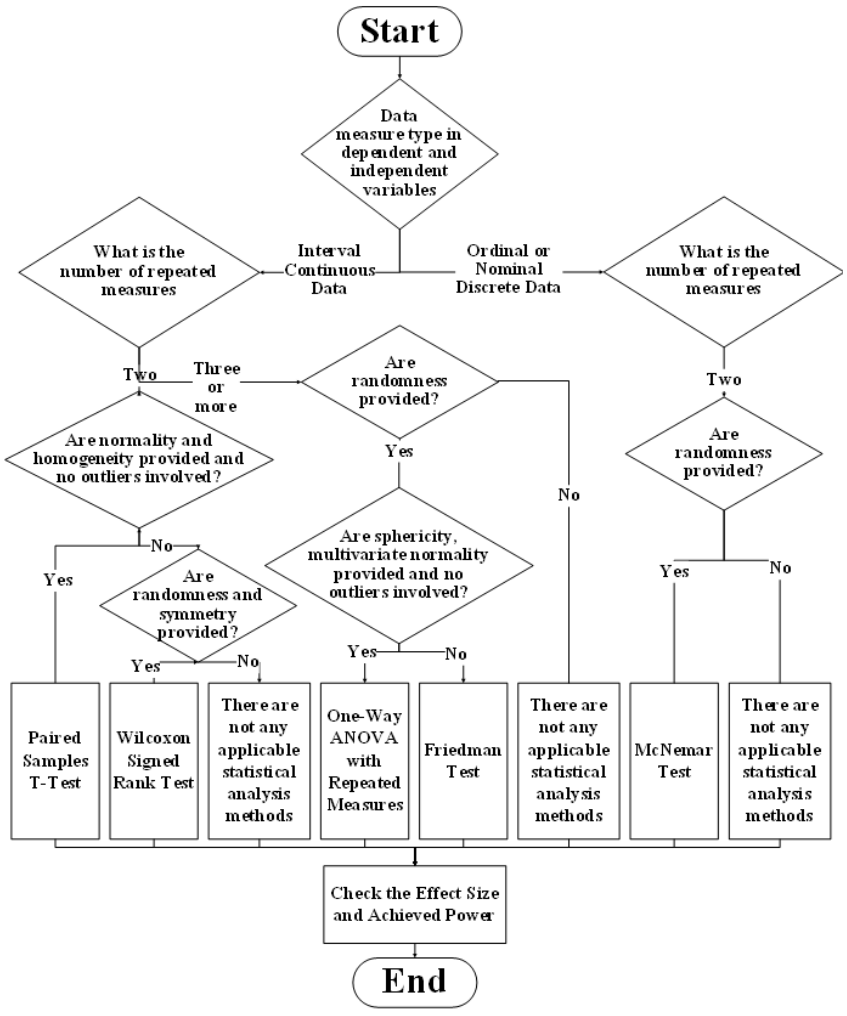


Figure 3.18. The flow chart of data analysis method selection for paired samples experiments.

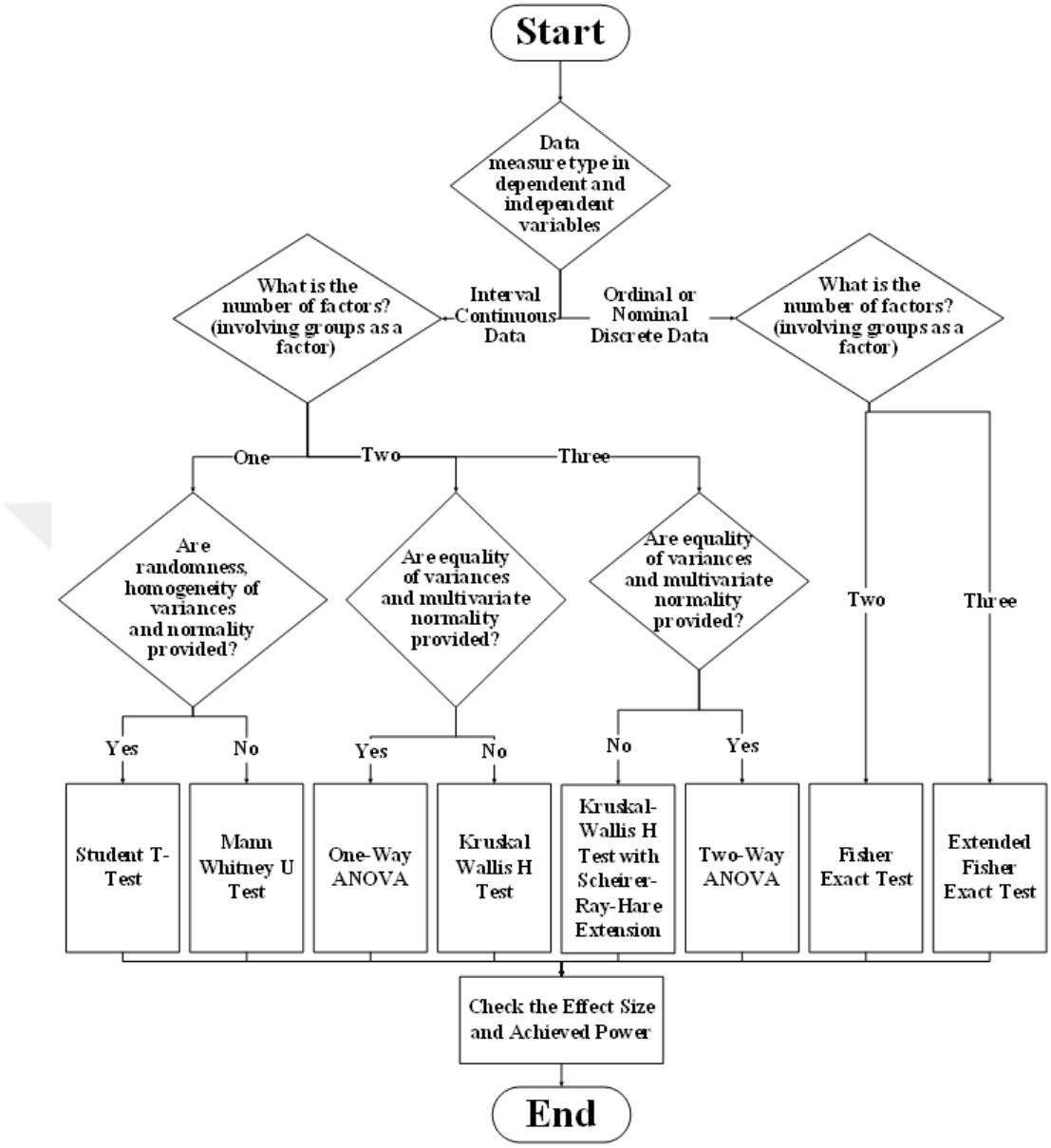


Figure 3.19. The flow chart of data analysis method selection for independent samples experiments.

3.4. Experimental Design

This section describes the conducted experiments to evaluate the developed hypotheses explained in the previous part of the research.

3.4.1. Information Comprehensibility Experiment

The methodological approach was to quantitatively analyze the effectiveness of V-SAFE based on mnemonics and understandability in a six-step experiment procedure. The experiment is conducted through three trainees on each session, due to the number of roles assigned in the V-SAFE. As the initial step, three roles, which are crane operator, pointer and bricklayer, assigned to three trainees in the session. The experiment was conducted through six groups; in other words, eighteen trainees enter the experiment. All trainees entered the experiment, are graduate students from Boğaziçi University Civil Engineering Department participated in the experiments. These students continue their graduate studies in the field of civil engineering, and they completed their on-site internship.

As the second step, of the experiment, all trainees watched a recorded PowerPoint presentation, which involves the essential safety information and regulations. The video recorded PowerPoint presentations were provided to the trainees based on each different role. For instance, if a trainee was assigned as a bricklayer role in the first stage, then he or she answered the questions about the bricklayer role. Based on common responsibilities and tasks defined in V-SAFE simulation, some questions of different roles were the same as each other. The video-recorded PowerPoint presentations were provided to the trainees for each different role, to ensure an unbiased and objective information gathering process. For instance, if lectures will be provided to the trainees by an instructor, rather than video-recorded PowerPoint presentation, potential forgotten information by the instructor in any session leads to a lack of information for the trainees. As a result trainees could not learn the essential information about the safety, and the potential error of an instructor could lead to inequality during the following stages.

In the third stage, all trainees entered to a paper-based examination with close-ended questions. This stage aims to evaluate the trainees' comprehension level about the lecture-based PowerPoint presentation. At this stage, trainees' comprehension level was assessed through their score on this close-ended examination. Trainees, who

will enter to V-SAFE training in the latter stages of the experiment, answered their roles' questions. To understand the impact of V-SAFE based on understandability and mnemonics, all questions are assigned to one of these factors.

As the fourth stage, all trainees watched a lecture-based video-recorded V-SAFE tutorial, which involves a tutorial about the usage of V-SAFE software. Also, all trainees obtained an infographic about the usage of V-SAFE. The infographics contain the necessary information required to use V-SAFE such as, how to select roles, controller configuration, simulation settings, etc. As a result, it was ensured that all the trainees received the same information based on their role, which was crucial to sustaining a controlled environment for the experiments.

As the fifth stage, all trainees entered to V-SAFE training. During the tower crane simulation, trainees were required to complete both personal and collaborative responsibilities in the virtual environment.

As the final stage, the answer sheets are returned to each trainee, and it is asked to revise their answers. The main aim of this stage was to monitor, whether or not trainees updated their answers', based on the training by V-SAFE. As a result, the evaluation of V-SAFE will be possible by checking the changes in the responses of the trainees.

After the completion of the data gathering process through V-SAFE training, the gathered data from the subjects were analyzed by using several statistical analysis methods. Firstly, the distribution of the dataset was analyzed through the Shapiro-Wilk test. Secondly, the homogeneity of the distribution was analyzed through the Levene test of homogeneity. Thirdly, Wald-Wolfowitz Runs test was conducted to evaluate, whether the randomness is provided in the dataset. Moreover, Cohen t-test was conducted to analyze the effect size and post-hoc power to check, whether the required sample size is achieved. Finally, based on the results of normality and homogeneity tests, Wilcoxon-Signed Rank Test was conducted. Secondly, each question on the dataset is assigned to the mnemonics and understandability features. The

randomness and achieved power were evaluated. Finally McNemar test was conducted to analyze the improvement on mnemonics and understandability. Table 3.3 illustrates the experimental setting of the information comprehensibility experiment.

Table 3.3. Experimental setting of the information comprehensibility experiment.

Stage 1 - Assigning Each Trainees' Roles and Groups
Instrucktor assigns each trainees' role in the session for the following stages of the experiment Objectives:Assigning each trainees' role, in order to provide them relevant V-SAFE tutorial, and necessary infographic in the latter stages of the experiment
Stage 2 - Conventional Safety Training
Trainees watched a video-recorded PowerPoint presentation that involves the neccestities of on-site safety Objective: Trainees learn the Essentials of the safety context
Stage 3 - First Examination
Trainees enter into close-ended examination Objective: Evaluating the trainees' comprehension level about the lecture-based PowerPoint presentation.
Stage 4 - V-SAFE Tutorial and Infograhics
Trainees watch a video-recorded PowerPoint presentation and obtain an infographic about V-SAFE controller configuration Objective: Trainees learn the Essentials of the safety context about their specified roles assigned in the previous stage and gather relevant information about the controllers of V-SAFE, thus, they could select the preffered input device
Stage 5 - V-SAFE Training
Trainees enter into V-SAFE virtual training Objective: Trainees could improve their comprehension level about the safety context through visual elements and tasks defined in V-SAFE
Stage 6 - Second Examination
Trainees enter into same close-ended examination Objective: Previous answer sheets in the stage 3 are delivered to the trainees and asked them to revişe their answers.
Stage 7 - Data Analysis
Gathered data is evaluated through several analysis method. Objective: Homogeneity, normality and randomness of the data is analyzed. Power analysis is conducted. Finally, non parametric paired t-test was conducted. Secondly, questions are categorized based on mnemonics and understandability. The randomness of the data is analyzed. The power analysis were conducted. Finally, McNemar test is conducted

3.4.2. Repeatable Virtual Experiment

The methodological approach was to quantitatively analyze the effectiveness of V-SAFE based on knowledge development and instructional feedback mechanism in a four-step experiment procedure. The experiment is conducted through three trainees on each session, due to the number of roles assigned in the V-SAFE. As the initial step, three roles were assigned to three trainees in the session. The experiment was conducted through six groups; in other words, eighteen trainees entered the experiment. All trainees entered the experiment, were undergraduate students from Boğaziçi University Civil Engineering Department participated in the experiments. These students continued their undergraduate studies in the field of civil engineering, and they completed their on-site internship.

As the second step of the experiment, all trainees watched a lecture-based video-recorded V-SAFE tutorial, which involves (i) a tutorial about the usage of V-SAFE software and (ii) the essentials of the safety behavior about the specified role. The video-recorded PowerPoint presentations were provided to the trainees for each different role, to ensure an unbiased and objective information gathering process. For instance, if lectures were presented to the trainees by an instructor, rather than video-recorded PowerPoint presentation, potential forgotten information by the instructor in any session, trainees could not have learned the essential information about the safety. Therefore, as also mentioned in information comprehensibility experiment settings, the error of an instructor could lead to a lack of information for trainees. Besides, all trainees obtained an infographic about the usage of V-SAFE. The infographics involved the necessary information required to use V-SAFE, such as how to select roles, controller configuration, simulation settings, etc. As a result, it was ensured that all the trainees received the same information based on their role, which was crucial to sustaining a controlled environment for the experiments.

As the third stage, all trainees enter to V-SAFE training. During the tower crane simulation, trainees are required to complete both personal and collaborative responsibilities in the virtual environment. The trainees' performance on V-SAFE is

automatically analyzed based on their behavior and decisions during the simulation. The safety performance of all trainees is captured, to make a comparative analysis between their performances on the latter stages. If all trainees completed both personal and collaborative responsibilities on the first trial, then their scores were counted as 1 on the latter stages. In this sense, unless simulation ends with flawless performance for all trainees, the experiment was continued for this group. However, if they failed on the previous stage, the trainees would re-enter to V-SAFE training second time and/or the third time.

After the completion of V-SAFE training, the gathered data from the subjects were analyzed by using several statistical analysis methods. Firstly, the multivariate normality distribution of the dataset was analyzed. Secondly, sphericity and randomness of the distribution were analyzed. Thirdly, Cohen t-test is conducted to examine the effect size, and power test was conducted to check, whether the required sample size. Finally, the One-Way ANOVA with Repeated Measures was conducted. Table 3.4.2 illustrates the experimental setting of the information comprehensibility experiment.

Table 3.4. Experimental setting of the repeatable virtual experiment.

Stage 1 - Assigning Each Trainees' Roles and Groups
Instructor assigns each trainees' role in the sessions for the following stages of the experiment. Objective: Assigning each trainees' role, in order to provide them relevant video-recorded lecture-based safety training, and necessary infographic in the latter stages of the experiment.
Stage 2 - V-SAFE Tutorial and Infograhics
Trainees watch a video-recorded PowerPoint presentation and obtain an infographic about V-SAFE controller configuration. Objective: Trainees learn the Essentials of the safety context about their specified roles assigned in the previous stage and gather relevant information about the controllers of V-SAFE, thus, they could select the preffered input device.
Stage 3 - First V-SAFE Training
Trainees enter into V-SAFE virtual training Objective: Traniees could learn the Essentials of the safety context, and improve their safety knowledge through hands-on-practice.
Stage 4 - Second V-SAFE Training
Trainees re-enter into V-SAFE virtual training Objective: Trainees could improve their safety knowledge through repetable practice.

Table 3.4. Experimental setting of the repeatable virtual experiment. (cont.).

Stage 5 - Third V-SAFE Training
Trainees re-enter into V-SAFE virtual training Objective: Trainees could improve their safety knowledge through repetable practice.
Stage 6 - Data Analysis
The safety performance of trainees automatically captured by V-SAFE in each stage (1st, 2nd and 3rd) were compared. Objective: The multivariate normality, sphericity and randomness of the data is analyzed. Moreover, the effect size is evaluated and the power analysis is conducted. Finally, based on the conducted descriptive analyses, ANOVA with repeated measures is conducted.

3.4.3. Hazard Identification Experiment

The methodological approach was to quantitatively analyze the effectiveness of V-SAFE based on hazard identification in a four-stage experiment procedure. All sessions are conducted through one trainee. The experiment is done through two groups, while, there are ten sessions on each group. In other words, twenty trainees enter the experiment. Twenty students from Boğaziçi University Civil Engineering Department participated in the experiments. These students continue their undergraduate and graduate studies in the field of civil engineering. The students who participated in the experiments completed their construction field internship for at least fifteen days.

As the second stage, trainees in the first group examined text-based safety notes, which involves the essential safety information and regulations. Thus, students in the first group could be able to identify each risk sources. Text-based information was provided to the first group trainees, to ensure an unbiased and objective information gathering process. At the same stage, trainees in the second group obtain an infographic about the usage of V-SAFE. The infographics involve the necessary information required to use V-SAFE such as, how to select roles, controller configuration, simulation settings, etc. Later on, trainees in the second group enter to V-SAFE virtual environment and monitor the risky situations and potential accident sources.

After the completion of the second stage, all trainees identified the potential risk sources. The hazard identification process is completed through, writing the potential safety hazard sources down in their own words. Rather than using a close-ended hazard identification list, the open-ended method is used to clarify the level of trainees' hazard identification. Close-ended methods such as checklists might support the trainees to orient or remember the hazard sources. For instance, when a trainee monitors fire and explosion as a risk source in the list, they could easily identify this risk source. Checking a close-ended list manipulates the results by orienting trainees to confirm all asked risk sources. However, when the question is open-ended, the trainees could not be able to identify a risk source if they do not learn anything about it. Also, trainees might not remember that risk source. As a result, open-ended questions found more suitable during the hazard identification process. All verbal and written statements of trainees matched with the relevant groups of risk sources. For instance, a trainee stated that crane operator might fall when he is climbing to the crane, this statement is marked as falling from a height.

As the third stage, the students in the first group toured the V-SAFE environment, and they revised the risk sources identified in the first stage. Therefore, they improve the number of recognized risk sources, which are defined during V-SAFE process. Similarly, the students in the second group examined written texts on occupational safety and were asked to revise the sources of risk encountered at the construction sites. Briefly, each group completed other groups' tasks in the previous stage.

The experiment setting does not allow the performance decrease in the second phase. In the first phase of the experiment, participants identify potential safety hazard sources based on text-based safety notes. In the second phase of the experiment, participants involved additional safety hazard sources in their original list after walking through the virtual environment. Therefore, technically, it is not possible to observe a lower risk identification level in the second part of the experiment. To cover this limitation, the analysis was conducted through two different methods. Firstly, the improvement in each group were analyzed through the McNemar test. As the initial descriptive analyses, randomness and power were also analyzed. Dataset of one groups'

pairs failed to provide randomness and power, which is crucial for the significance of the Fisher Exact Test. Therefore, the results were evaluated for one paired group. Secondly, it was aimed to assess the difference in each group through the Fisher Exact test, if dataset provides necessary assumptions. However, the dataset fails to provide essential assumptions of Fisher Exact Test such as randomness and power. Consequently, only the results of paired samples in one group were evaluated. Table 3.5 shows the steps of the experiment included in the study.

Table 3.5. Experimental setting of the hazard identification experiment.

Stage 1 - First Hazard Identification Training	
Conventional Hazard Identification Process	V-SAFE Hazard Identification Process
Trainees in the first group examined the text-based safety notes. Objectives: Trainees learn the essantials of the safety context about each hazard sources occured on the real construction sites.	Trainees in the second group obtain an infographic about V-SAFE controllers. Objective: Trainees could gather relevant information about the controllers of V-SAFE, thus, they could select the preffered input device. Trainees in the second group tour the virtual construciton site and monitör potential risk sources occuredin V-SAFE. Objective: Trainees could monitör virtual environment, thus, they could identify potential on-site risks.
Stage 2 - First Identification of the Hazard Sources	
All trainees identified the potential risk sources. Objective: Gathering the first part of necessary data to evaluate the effectiveness of both conventional and virtual methods during hazard identification process.	
Stage 3 - Second Hazard Identificaiton Training	
V-SAFE Hazard Identification Process Trainees in the first group obtain an infographic about V-SAFE controllers. Objective: Trainees could gather relevant information about the controllers of V-SAFE, thus, they could select the preffered input device. Trainees in the first group tour the virtual construction site and monitör potential risk sources occured in V-SAFE. Objective: Trainees could monitör virtual environment, thus, they could identify potential on-site risks.	Conventional Hazard Identification Process Trainees in the second group examinated the text-based safety notes. Objective: Trainees learn the Essentials of the safety context about each hazard sources occured on the real construction sites.
Stage 4 - Second Identification of the Hazard Sources	
All trainees identified the potential risk sources. Objective: Gathering the second part of necessary data to evaluate the effectiveness of both conventional and virtual methods during hazard identification process.	
Stage 5 - Data Analysis	
The difference between the improvement in the groups is evaluated through several analysis methods Objective: The randomness of the data is analyzed. The power analysis were conducted. Finally, McNemar testis conducted.	

3.4.4. Eye-Movement Behavior Comparison Experiment

To determine whether V-SAFE provides an effective hands-on practice, a four-staged experiment procedure is prepared. Eleven construction workers from Bertuğ Bey Construction participated in the experiment. In each session, one worker enters the real construction site. Later on, the same worker enters the V-SAFE virtual environment. During touring the actual construction site, it is asked workers to identify potential hazards. The total fixation duration, first fixation duration, time to the first fixation, and fixation count during the experiment were recorded for further analysis. Tobii ProGlasses 2 eye-tracker was used to record each workers' visual attention movement, including their fixation, visit, and saccade on the real construction site (Figure 3.20). Each worker interacted with the instructor and identified the potential hazards by their verbal statements, and the instructor interviewed the workers, following the experiment.



Figure 3.20. Gathering eye-tracking metrics through Tobii X2-30 and Tobii Pro Glasses 2.

In the second stage, workers watch a lecture-based video-recorded PowerPoint presentation, which involves a tutorial about the usage of V-SAFE software. The video-recorded PowerPoint presentations were provided to the trainees to ensure an unbiased and objective information gathering process. In addition, all workers obtain an infographic about the usage of V-SAFE. The infographics involve the necessary information required to use V-SAFE such as controller configuration, simulation settings, etc.

As the third stage, the same construction workers in the first stage entered the V-SAFE virtual environment. Before the interaction with V-SAFE, the experiment facilitator supported each worker in calibrating the Tobii X2-30 on-screen eye-tracker. Subsequently, the workers commenced the experiment by inspecting the virtual construction site to identify potential hazards without any time limitations. Workers identify the hazards by verbal notice.

After the completion of the training, potential risk sources were defined as the area of interests (AOI) on the real construction site and V-SAFE. For instance, potential risk sources such as ladders without a railing, nails on the construction site, which exist on the viewpoint of the trainees is defined on the real construction site and V-SAFE (Figure 3.21).

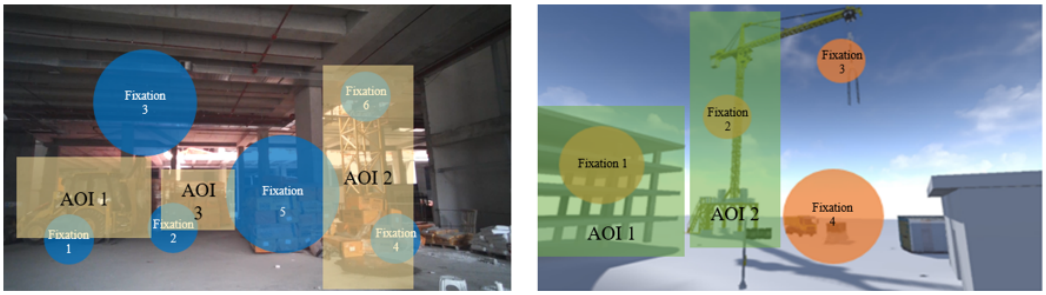


Figure 3.21. The sample assignment of the area of interests on real construction site and V - SAFE.

Consequently, the necessary eye-tracking data: (i) total fixation duration, (ii) first fixation duration, (iii) time to first fixation and (iv) fixation count gathered. Firstly, the normality distribution is analyzed through the Kolmogorov-Smirnov test. Secondly, the homogeneity of the distribution is analyzed through the Levene test of homogeneity. Thirdly, Cohen t-test is conducted to analyze the effect size, and power test is conducted to find out the required sample size. Moreover, Wald-Wolfowitz Runs Test conducted to check, whether the sample group is random from the population. Finally, the nonparametric Mann-Whitney U test is performed. Table 3.6 shows the steps of the experiment included in the study.

Table 3.6. Experimental setting of the eye-movement behavior comparison experiment.

Stage 1 - Eye Tracking Calibration
Workers put on mobile eye-tracker device. Experiment facilitator supported trainees in calibrating the eye-tracker.
Stage 2 - Real Construciton Site Hazard Identification
Trainees enter to real construction site to identify on-site risks Objectives: -To gather eye tracking data to the trainees -To provide identified hazard sources by the trainees
Stage 3 V-SAFE Tutorial and Infograhics
Workers watch a video-recorded PowerPoint presentation and obtain an infographic about V-SAFE controller configuration. Objectives: Trainees learn the Essentials of the safety context about their specified roles assigned in the previous stage and gather relevant information about the controllers of V-SAFE, thus, they could select the preffered input device.
Stage 4 - V-SAFE Hazard Identification
Trainees enter into V-SAFE environment. Objectives: -To gather eye tracking data of the trainees -To provide identified hazard sources by the trainees
Stage 5 - Data Analysis
The eye tracking data of the trainees will be analyzed through comparative analysis methods. Objectives: To understand the difference between the trainees eye tracking behavior between real site interaction and V-SAFE interaction, several analyses were conducted. Firstly, the homogeneity, normality, randomness of the data is alayzed. Secondly, effect size and power analysis were conducted. Finally, non parametric student t-test were conducted.

3.4.5. Construction Site and Gaming Experience Experiment

In order to assess the impact of the virtual work safety training tool on trainees without any construction site and gaming experience, two different study groups were compared. All trainees in each group are either graduate or undergraduate students from various universities. The first group was composed of a total of twelve students with a civil engineering background. Besides, subjects in Group 1 completed site internship, and they are familiar with construction site activities. The second group was composed of twelve students from other disciplines such as computer engineering, mechanical engineering, etc. Subjects in Group 2 do not have any construction experience about the construction site environment.

It is developed a four-staged experiment procedure. In the first stage, all trainees watched a video-recorded PowerPoint presentation for their specified role. The recorded PowerPoint presentation provides (i) a tutorial about the usage of V-SAFE software and (ii) the essentials of the safety behavior about the specified role. Moreover, the trainees obtained an infographic that involves the necessary information about the controller configuration of V-SAFE. Furthermore, it was asked all the trainees to select their preferred input device. Some trainees found it challenging to use the keyboard and mouse layout, so they chose gamepads, which provide them a simpler interface.

In the second stage, all trainees enter V-SAFE training. During the tower crane simulation, trainees are required to complete both personal and collaborative responsibilities in the virtual environment. The trainees' performance on V-SAFE is automatically analyzed based on their behavior and decisions during the simulation. The safety performance of all trainees' is captured, to make a comparative analysis between their performances on the latter stages.

In the third stage, all trainees received specific feedback from V-SAFE based on their safety behavior during the virtual training. After the completion of V-SAFE training, a survey was conducted on a Likert scale to determine the participants' gaming experience level. Accordingly, we gathered the necessary data and examined the correlation between the gameplay experience of the trainees and their virtual safety training score.

After the completion of V-SAFE training, the gathered data from the subjects were analyzed by using several statistical analysis methods. Firstly, the normality of distribution was analyzed. Secondly, equality of variances was analyzed. Thirdly, Cohen t-test was conducted to analyze the effect size, and power test is conducted. Finally, based on the results of descriptive analyses, Scheirer-Ray-Hare Test was performed. Table 3.7 illustrates the experimental setting of the construction site and gaming experience experiment.

Table 3.7. Experimental setting of the construction site and gaming experience experiment.

Stage 1 - V-SAFE Tutorial and Infograhics
Trainees watch a video-recorded PowerPoint presentation and obtain an infographic about V-SAFE controller configuration. Objectives: Trainees learn the Essentials of the safety context about their specified roles and gather relevant information about the controllers of V-SAFE, thus, they could select the preffered input device.
Stage 2 - V-SAFE Training
Trainees enter into V-SAFE virtual training. Objectives: Trainees could improve their comprehension level about the safety context through visual elements and tasks defined in V-SAFE
Stage 3 - Filling in Questionnaires
Trainees fill in the questionaires at the end of V-SAFE simulation Objectives: -Trainees inform their demographic information through survey (expertise, age, gaming experience). -Trainees fill in questionnaires to Express their opininons about V-SAFE training.
Stage 4 - Data Analysis
The difference between the 1st and 2nd groups is evaluated through several analysis method Objectives: The normality, equality of varianhces of the data were analyzed. Moreover, the effect size is evaluated and the power analysis is conducted. Finally, based on the conducted descriptive analyses, Scheirer-Ray-Hare Test is executed

4. ANALYSIS RESULTS

This chapter presents a statistical analysis of the collected data. The experiment population for the data collection was the undergraduate, graduate students from the various universities and construction workers from several construction companies. A total of five experiment results are represented. Besides, the statistical findings are summarized, and the descriptive analyses are presented.

4.1. Information Comprehensibility Experiment

This experiment is conducted to evaluate the following hypotheses:

- Hypothesis 1.1: V-SAFE provides effective demonstration process for the trainees, thus, significantly improves the understandability of safety information, when used after PowerPoint presentations.
- Hypothesis 1.2: V-SAFE significantly improves the spatial mnemonics of trainees when used after PowerPoint presentations

The results of each trainee are illustrated in Table 4.1 and Table 4.2, the assignment of each question to mnemonics and understandability is shown in Table 4.3.

Table 4.1. Information Comprehensibility Experiment Results.

		PowerPoint-based lectures									V-SAFE								
Trainee	**Role**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
1	Crane Operator	1	0	1	0	1	0	0	1	0.500	1	0	1	1	1	0	1	1	0.750
2	Crane Operator	1	0	0	0	1	0	0	0	0.250	1	0	1	0	1	0	0	0	0.375
3	Crane Operator	0	0	0	1	1	0	0	0	0.250	1	0	1	1	1	1	0	1	0.750
4	Crane Operator	1	0	1	1	1	0	0	1	0.625	1	0	1	1	1	0	1	1	0.750
5	Crane Operator	1	0	1	1	1	0	0	1	0.625	1	0	1	1	1	0	0	1	0.625
6	Crane Operator	1	0	1	1	1	1	0	0	0.625	1	0	1	1	1	1	0	1	0.750
7	Pointer	1	0	1	1	1	0	0	0	0.5	1	0	1	1	1	0	1	0	0.625

Table 4.1. Information Comprehensibility Experiment Results (cont.).

		PowerPoint-based lectures									V-SAFE								
Trainee	**Role**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
8	Pointer	1	0	1	1	0	0	0	0	0.375	1	1	1	1	0	0	0	0	0.500
9	Pointer	1	0	0	1	0	1	0	0	0.375	1	0	0	1	1	0	0	0	0.375
10	Pointer	1	0	1	0	1	0	1	0	0.500	1	0	1	1	1	0	1	0	0.750
11	Pointer	1	0	1	1	0	0	0	0	0.375	1	1	1	1	1	1	0	0	0.750
12	Pointer	0	0	1	0	0	1	0	0	0.250	0	0	1	0	0	1	0	0	0.250
13	Brick Layer	1	0	1	0	0	0	1	1	0.500	1	1	1	0	0	1	1	1	0.750
14	Brick Layer	1	1	1	1	0	0	1	0	0.625	1	1	1	1	1	0	1	0	0.750
15	Brick Layer	1	1	1	0	1	0	1	1	0.750	1	1	1	0	1	0	1	1	0.750
16	Brick Layer	1	1	1	0	0	0	1	1	0.625	1	1	1	1	0	0	1	1	0.750
17	Brick Layer	1	1	1	1	1	0	1	1	0.875	1	1	1	1	1	0	1	1	0.875
18	Brick Layer	1	0	1	1	1	1	1	0	0.75	1	1	1	1	1	1	1	0	0.875
All Trainees:										0.524	All Trainees:								0.660
In the table, 1 denotes correct answers, while 0 denotes incorrect answers.																			

Table 4.2. Information Comprehensibility Experiment Results based on Information Quality Parameters

Trainee	**Questions Related to Mnemonics**		**Questions Related to Understandability**	
	PowerPoint-based Lectures	**V-SAFE**	**PowerPoint-based Lectures**	**V-SAFE**
1	0.75	1.00	0.00	0.00
2	0.50	0.75	0.00	0.00
3	0.50	1.00	0.00	1.00
4	1.00	1.00	0.00	1.00
5	1.00	1.00	0.00	0.33
6	1.00	1.00	0.33	0.33
7	1.00	1.00	0.25	0.75
8	0.67	0.67	0.25	0.75
9	0.33	0.67	0.50	0.75
10	1.00	1.00	0.25	0.25
11	0.67	1.00	0.25	0.25
12	0.33	0.33	0.25	0.50
13	0.50	0.75	0.50	1.00
14	0.50	0.75	0.75	0.50
15	0.75	0.75	0.75	0.50
16	0.50	0.50	0.75	0.50
17	0.75	0.75	1.00	0.25
18	1.00	1.00	0.50	0.75
Overall	0.7083	0.8287	0.3519	0.5231

Table 4.3. Categorization of Questions.

Role	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Crane Operator	M	U	M	M	M	U	U	U
Pointer	M	U	M	U	M	U	U	U
Brick Layer	M	U	U	U	M	M	U	U
M denotes the mnemonics, U denotes understandability								

The average of correct answers in the first multiple-choice examination was 0.524, while, the average of correct answers in the second multiple-choice examination was 0.660 (Figure 4.1). This corresponds to an average increase of 0.136. Also, when all questions were linked to the information quality parameters as mnemonics and understandability (Figure 4.2), there was still a considerable improvement. To illustrate, the trainees' performance in the questions related to mnemonics were 0.7083, when they only watched PowerPoint lectures. After the completion of V-SAFE training, their score improved to 0.8287, which corresponds to 0.1204 improvements. The same situation exists on the questions related to understandability. The trainees' examination score before and after V-SAFE training was 0.3519 and 0.5231 respectively, which corresponds to 0.1712. In the first overview, there is a potential improvement determined in the dataset. It was noted that the dataset involves nominal data. When analyzing an ordinal binary dataset, parametric approach, such as paired t-test is not suitable, since binary data is usually skewed, in other words, datasets involve many smaller values, few higher values (McElduff *et al.*, 2010). To evaluate the skewed data, the previous literature emphasizes that using a non-parametric paired t-test, namely Wilcoxon signed-rank test is a more suitable approach (McElduff *et al.*, 2010). However, as an essential assumption of Wilcoxon signed-rank test, the dataset used in the analysis should contain continuous data (Wilcoxon *et al.*, 1970). To evaluate a nominal binary data in paired groups, McNemar method is a suitable method. The main assumptions of McNemar test involve one ordinal variable with two categories and one independent variable with two connected groups (McCrum-Gardner, 2008). To select the appropriate data analysis method, it is crucial to define the experimental setting of

the conducted experiment and data characteristics (Figure 4.4). According to dataset in the experiment, McNemar Test is selected as the most appropriate method (Figure 4.1).

Table 4.4. Experiment and Data Features for Paired Samples.

Experiment	**Safety Information Transfer**
Data Measure Type of Dependent Variable and Independent Variables	Categorical, Ordinal Data
Data Characteristics	Binary, Discrete Data
Experimental Procedure	Paired Samples
Number of Repeated Measures	2
Potential Analysis Method	McNemar Test

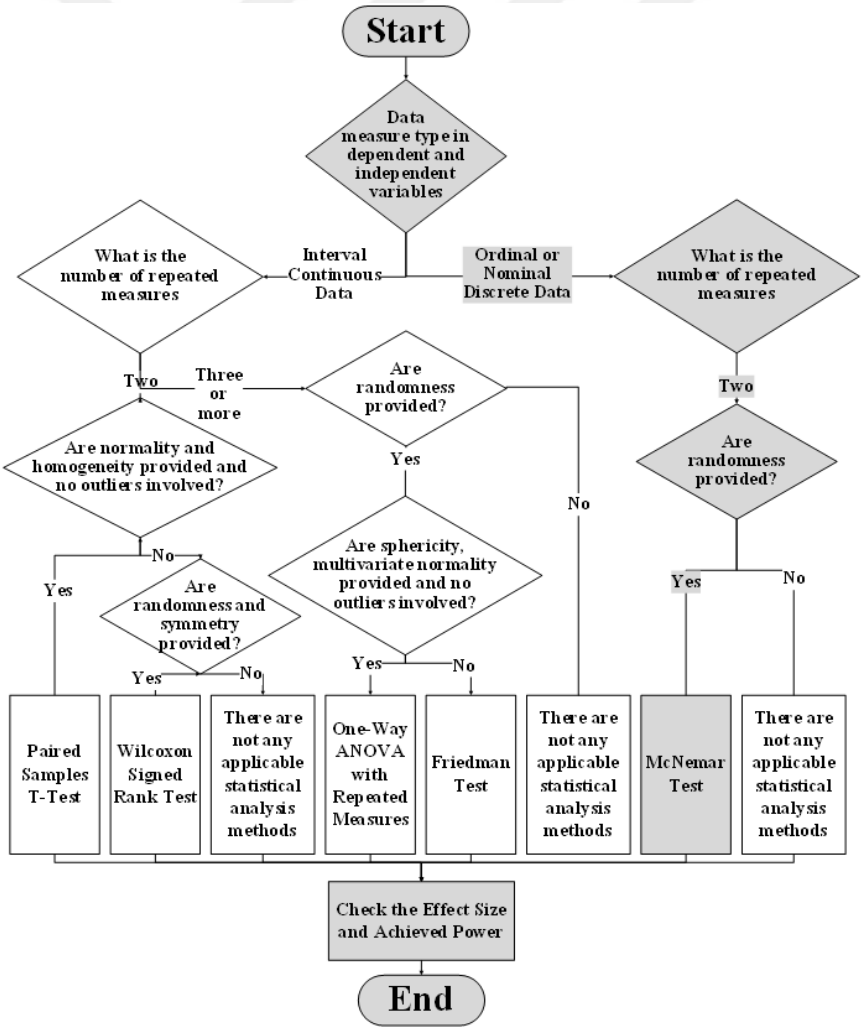


Figure 4.1. The Appropriate Data Analysis Method for Information Comprehensibility Experiment.

4.1.1. Descriptive Analyses

According to the previous literature, the significant assumptions of the McNemar test is randomness (Riffenburgh, 2006). In addition, it is crucial the evaluate power, to ensure the required sample space was met. As the preliminary analysis of data, randomness test and power tests were conducted.

4.1.1.1. Testing the Power. The achieved power is an important parameter that must be taken into consideration in McNemar Test (Kılıç, 2014). To analyze the effect size, power analysis is performed. The analysis results are illustrated in Figure 4.5.

Table 4.5. The Summary of the Power Test.

Risk	**Trial**	**Odds Ratio**	**Sample Size**	**The Proportion of Discordant Pairs**	**Lower Critical N**	**Upper Critical N**	**Actual α**	**Power (1-β)**
Mnemonics	Lecture Notes V-SAFE	2.06	120	0.43	33	33	0.0352	0.7729
Understandability	Lecture Notes V-SAFE	441	17	0.03	32	32	0.046	0.8255

4.1.1.2. Testing the Randomness. On of the major assumptions of the McNemar test is the randomness (Taeger and Kuhnt, 2014). To check, whether or not randomness is provided in the data, Wald-Wolfowitz Runs Test is performed. The analysis results are illustrated in Table 4.6.

Table 4.6. The summary of randomness test analysis results

	Mnemonics		Understandability	
	PowerPoint Lecture	V-SAFE	PowerPoint Lecture	V-SAFE
Runs	25	21	33	39
N0	19	11	50	38
N1	41	49	34	46
N	60	60	84	84
Expected Runs	29.967	18.967	41.476	42.619
Variance Runs	10.988	5.167	19.251	20.368
Standard Deviation Runs	-3.315	2.273	4.388	4.513
Z	-0.593	0.895	-1.932	-0.802
ρ-Value	0.276	0.814	0.053	0.423

Analysis results highlight that the datasets provided the randomness. In addition, the achieved power of dataset related to mnemonics is 0.7729 (Table 4.5). In default, the value should be 0.80 as default (Ellis, 2010). However, because the difference is quite low, the power value is acceptable in the dataset related to mnemonics. In the dataset related to understandability, achieved power is 0.8255, which is more than 0.80. Therefore, it is suitable to conduct the McNemar test.

4.1.2. Analysis Results

In statistics, the McNemar is a statistical test used in paired binomial data. Due to the dataset of each risk source is binomial, it is determined to use the McNemar test to evaluate whether the row and column marginal frequencies are equal. The analysis results of the McNemar test is illustrated in Table 4.7. Consequently, the status of the hypotheses is represented as following (Table 4.8).

Table 4.7. Wilcoxon Signed-Rank Test Results.

Related Information Quality Parameter	**Trial**	**Mean**	**St. Dev.**	**Sample Size**	**Exact Significance**	**Impact**
Mnemonics	Lecture	0.683	0.469	60	0.031	**
	V-SAFE	0.816	0.39	60		
Understandability	Lecture	0.405	0.494	17	0.001	***
	V-SAFE	0.548	0.501	17		
In the first group. trainees firstly identified hazard sources through V-SAFE. Then, they identified hazard sources through lecture notes. Analysis results: *: $\rho < 0.10$. **: $\rho < 0.05$. ***: $\rho < 0.01$. N.S.= not statistically significant.						

Table 4.8. The status of the hypotheses.

Hypotheses	**Abbreviation**	**Status**
V-SAFE provides effective demonstration process for the trainees, thus, significantly improves the understandability of safety information, when used after PowerPoint presentations.	H1.1.	Accepted ($\rho < 0.05$)
V-SAFE significantly improves the spatial mnemonics of trainees when used after PowerPoint presentations	H1.2.	Accepted ($\rho < 0.01$)

4.2. Repeatable Virtual Experiment

This experiment is conducted to evaluate the following hypotheses:

- Hypothesis 2.1: Trainees using V-SAFE significantly correct their mistakes by repeatable practice.
- Hypothesis 2.2: V-SAFE significantly improves the implicit knowledge level of trainees by providing repeatable hands-on practice.

To evaluate the effectiveness level of feedback, the improvement in each trial are analyzed. The results of each trainee are illustrated in Table 4.9.

Table 4.9. Analysis results of the information comprehensibility experiment.

	Automated Safety Training Performance		
Trainee	**1st Trial**	**2nd Trial**	**3rd Trial**
1	1.0000	1.0000	1.0000
2	0.8500	1.0000	1.0000
3	0.8500	1.0000	1.0000
4	0.9500	1.0000	1.0000
5	0.7300	0.7800	0.9300
6	1.0000	1.0000	1.0000
7	0.8000	0.8000	1.0000
8	0.8000	1.0000	1.0000
9	1.0000	1.0000	1.0000
10	0.8000	1.0000	1.0000
11	0.9400	0.9400	0.9500
12	1.0000	1.0000	1.0000
13	0.7200	0.9000	1.0000
14	1.0000	1.0000	1.0000
15	1.0000	1.0000	1.0000
16	0.6700	1.0000	1.0000
17	0.9200	1.0000	1.0000
18	1.0000	1.0000	1.0000
Average	0.8906	0.9678	0.9933
Standard Deviation	0.1143	0.0700	0.0197
Sample Size	18	18	18

From the first overview, it was noticed that the increase in the automated safety training performances between 1st to 3rd trial was 0.1027, 1st to 2nd trial was 0.0772

and 2nd to 3rd trial was 0.0255. These increases point out a high potential for statistically significant improvement. Just like the information comprehensibility experiment described in the previous section, this experiment is also conducted through the same trainees in a paired experimental design. The data of each trainees' results are an interval, continuous and the number of repeated measures are three. To analyze the differences between each group that contain three interval continuous data, two data analysis methods as One-Way ANOVA with Repeated Measures and Friedman tests existed in the literature (Table 4.15). To select the appropriate data analysis method, it is essential to define the experimental setting of the conducted experiment and data characteristics (Table 4.10).

Table 4.10. Experiment and Data Features.

Experiment	Repeatable Virtual Experiment
Data Measure Type of Dependent Variable and Independent Variables	Interval Data
Data Characteristics	Continuous Data
Experimental Procedure	Paired Samples
Number of Repeated Measures	3
Potential Analysis Methods	One-Way ANOVA with Repeated Measures and Friedman test

4.2.1. Descriptive Analyses

As the preliminary analysis of data, several tests were conducted to understand the features and characteristics of the data. The lack of sphericity of covariance across repeated measures enlarges the error rates of type I in the univariate repeated-measures analysis of variance (Muller and Barton, 1989). Without sphericity, there are not any methods available to compute power (Muller and Barton, 1989) and the impact of violating sphericity causes loss of power, and that results from the trivial comparison of F-distribution due to tabulated values (Field, 2013).

4.2.1.1. Testing the Sphecirity of the Data. To check the data, whether or not the sphericity is provided, and Mauchly test of sphericity is performed (Table 4.11).

Table 4.11. Mauchly Test of Sphericity Results.

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Time	0.629	7.406	2	0.025	0.73	0.781	0.5

According to the multivariate Mauchly test of sphericity results, Mauchlys' W equals to 0.629, which corresponds to ρ-value of 0.025. Therefore, the sphericity of the data is significantly provided in the dataset.

4.2.1.2. Testing the Effect Size, Power and Multivariate Normality. Several standardized measures exist, which summarize the strength of the correlation among the independent variables and dependent variable in the One-Way ANOVA with Repeated Measures model (Wilkinson, 1999). Standardized effect-size is one of these measures, which facilitates statistical comparison and not commonly used in One-Way ANOVA with Repeated Measures analysis (Wilkinson, 1999).

Another critical analysis, which is applied in the One-Way ANOVA with Repeated Measures is power analysis method. The power analysis method is executed to determine the probability of type II error, thus, evaluating the effect size in population, sample size, and significance level becomes possible (Howell, 2009). Accordingly, the multivariate test is performed to analyze the effect size, achieved power and multivariate normality. The analysis results are illustrated in Table 4.12.

Table 4.12. Multivariate Test Results.

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared	Noncentrality Parameter	Observed Power
Time	Pillai's Trace	0.474	7.206	2	16	0.006	0.474	14.411	0.879
	Wilks' Lambda	0.526	7.206	2	16	0.006	0.474	14.411	0.879
	Hotelling's Trace	0.901	7.206	2	16	0.006	0.474	14.411	0.879
	Roy's Largest Root	0.901	7.206	2	16	0.006	0.474	14.411	0.879

According to multivariate test results, observed power is 0.879, which highlights that the probability of type II error is higher than 0.8, which is highly acceptable in default. Besides, there is a large effect size between the groups. Moreover, multivariate test results show that partial eta squared value equals to 0.474. The effect size, which is 0.474 is significantly larger than 0.22 (see Table 3.2). As a result, the achieved effect size in the dataset is substantially larger than the required effect size. Finally, the significance value of Wilks' lambda (0.006) clearly shows that the dataset significantly provides multivariate normality.

4.2.1.3. Testing the Randomness. To check, whether or not randomness is provided in the data, Wald-Wolfowitz Runs Test is performed. The analysis results are illustrated in Table 4.13.

Table 4.13. The summary of randomness test analysis results.

	First Trial	Second Trial	Third Trial
Test Value	0.93	1	1
Cases < Test Value	9	4	2
Cases >= Test Value	9	14	16
Total Cases	18	18	18
Number of Runs	13	9	5
Z	1.215	0.924	0
Asymp. Sig. (2-tailed)	0.224	0.355	1

Based on each tested values on the Wald-Wolfowitz Runs Test, as could be seen from Table 4.12 that significance values are much higher than 0.05. Therefore, one can conclude that randomness is significantly provided on each trial in the dataset.

Descriptive analyses result show that the dataset significantly provides the assumptions of parametric One-Way ANOVA with Repeated Measures, which are multivariate normality, sphericity, and randomness. In this case, One-Way ANOVA with Repeated Measures should be conducted (Figure 4.2).

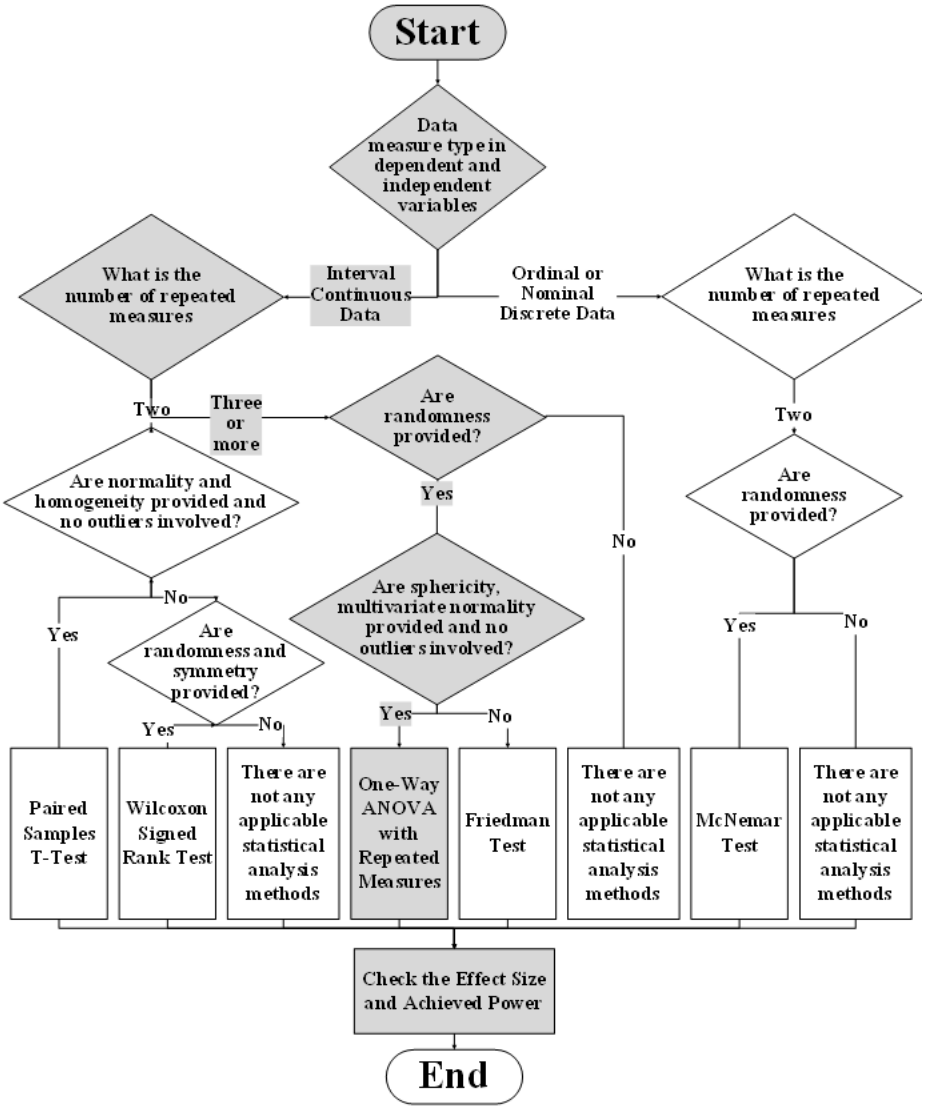


Figure 4.2. The Appropriate Data Analysis Method for Repeatable Virtual Experiment.

4.2.2. Analysis Results

According to the Wilks' Lambda values' significance equals to 0.006, which highlights a significant difference between at least one combinations of two different groups (e.g., 1st - 3rd Score, 1st - 2nd Score, 2nd - 3rd Score). To evaluate each groups' differences, posthoc analysis is performed (Table 4.14).

Table 4.14. Post-Hoc Analysis Results.

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Differenceb	
					Lower Bound	Upper Bound
1	2	-0.077*	0.024	0.013	-0.14	-0.015
	3	-0.103*	0.026	0.003	-0.173	-0.033
2	1	0.077*	0.024	0.013	0.015	0.14
	3	-0.026	0.014	0.263	-0.063	0.012
3	1	0.103*	0.026	0.003	0.033	0.173
	2	0.026	0.014	0.263	-0.012	0.063

Analysis results indicate that the significance of improvement between 1st and 3rd trials were 0.003, which is far less than 0.01. Moreover, the importance of growth among 1st and 2nd trials was 0.013, which is less than 0.05. However, the significance of the improvement between 2nd and 3rd trials was 0.263, which did not indicate a significant improvement. Therefore, the results indicate that the trainees' performance significantly improved between 1st and 2nd, and 1st and 3rd trials. However, there was not any significant improvement between 2nd and 3rd trial. Consequently, the status of the hypotheses is illustrated as following (Table 4.15).

Table 4.15. The status of hypotheses.

Hypothesis	Abbreviation	Status
Trainees using V-SAFE significantly correct their mistakes by repeatable practice	H.2.1	Accepted
V-SAFE significantly improves the implicit knowledge level of trainees by providing repeatable hands-on practice	H.2.2	Accepted

4.3. Hazard Identification Experiment

This experiment is conducted to evaluate Statement 3: *V-SAFE provides an effective hazard identification process for trainees.* Table 4.16 shows the hazard identification performance of the participants.

Table 4.16. Experiment Results of Hazard Identification Experiment.

Risk	1st Group: V - SAFE and Conventional Training			2nd Group: Conventional Training and V - SAFE		
	V-SAFE	Conventional Method	Improvement	Conventional Method	V-SAFE	Improvement
Fall from height	9	9	0	11	16	5
Electrocution	10	10	0	5	16	11
Building / soil collapse	0	0	0	5	8	3
Accident by falling object	2	2	0	8	16	8
Heavy equipment accidents	7	7	0	3	14	11
Fire or explosion	9	10	1	5	16	11
Traffic accidents on site	4	7	3	3	9	6
Materials-based risks	10	10	0	5	17	12
Total	**51**	**55**	**4**	**45**	**115**	**70**
Average	**5.1**	**5.5**	**0.4**	**2.647**	**6.764**	**4.117**

In the first group, trainees' average hazard identification performance was 5.100 in the V-SAFE environment, while their average was 5.500 after the completion of conventional safety training. This increase corresponds to an average of 0.400. In the second group, trainees identified an average of 2.647 hazards by using the conventional safety training method. After completion of V-SAFE training, their hazard identification average was 6.764, which corresponds to 4.117 improvement. In the first overview, these increases highlight the great potential for statistically significant growth. It was noted that the dataset contains binary data for each hazard sources, while, the total number of identified hazard sources are counted through the summation of identified hazard sources. In the experimental setting, two different evaluations existed.

Firstly, it is aimed to evaluate the trainees' improvement between conventional safety training method and V-SAFE for each hazard sources. Therefore, the experimental setting of the first analysis is conducted through the same trainees in a paired analysis procedure. The dataset involves ordinal binary data. When analyzing an ordinal binary data, parametric approach, such as paired t-test is not suitable, since

binary data is usually skewed, in other words, datasets involve many smaller values, few higher values (McElduff *et al.*, 2010). To evaluate the skewed data, the previous literature emphasizes that using a non-parametric paired t-test, namely Wilcoxon signed-rank test is a more suitable approach (McElduff *et al.*, 2010). However, as an essential assumption of the Wilcoxon signed-rank test, the dataset used in the analysis should contain continuous data (Wilcoxon *et al.*, 1970). In order to evaluate a ordinal binary data in paired groups, McNemar method is a suitable method. The main assumptions of McNemar test involve one ordinal variable with two categories and one independent variable with two connected groups (McCrum-Gardner, 2008). To select the appropriate data analysis method, it is crucial to define the experimental setting of the conducted experiment and data characteristics (Table 4.17). According to dataset in the experiment, McNemar Test is selected as the most appropriate method (Figure 4.3).

Table 4.17. Experiment and Data Features for Paired Samples.

Experiment	**Hazard Identification Experiment (Paired Samples)**
Data Measure Type of Dependent Variable and Independent Variables	Categorical, Ordinal Data
Data Characteristics	Binary, Discrete Data
Experimental Procedure	Paired Samples
Number of Repeated Measures	2
Potential Analysis Method	McNemar Test

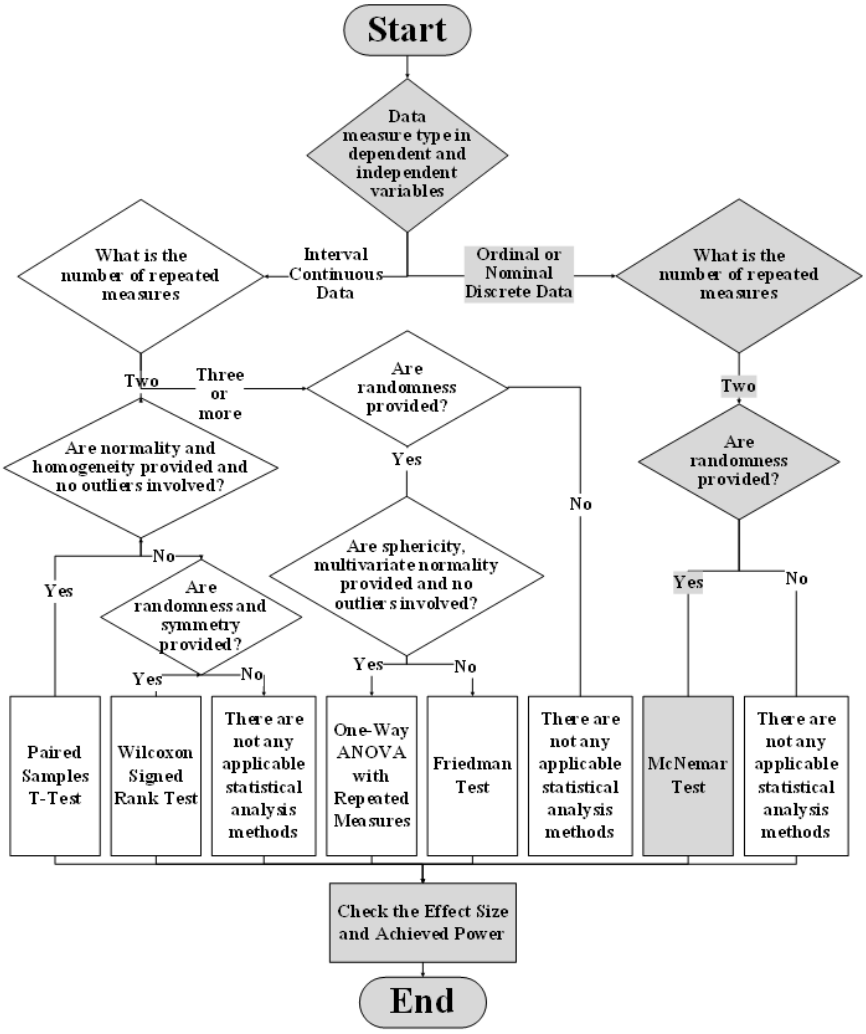


Figure 4.3. The Appropriate Data Analysis Method for Hazard Identification Experiment for Paired Differences.

Secondly, it is aimed to evaluate the trainees' improvement between two mutually exclusive groups for each hazard sources. Therefore, the experimental setting of the second analysis was conducted through trainees in two different groups. The dataset involves ordinal binary data. When analyzing an ordinal binary data, parametric approach, such as student t-test, Mann Whitney-U test, are not suitable. As an essential assumption of these tests, the dataset should contain continuous data (Wilcoxon *et al.*, 1970). Besides, the previous literature (McCrum-Gardner, 2008) states that if the sample size is less than 20, using Fisher's exact test is more suitable compared to the chi-square test. To evaluate binary nominal data, the Fisher exact test is selected. Fisher exact test is a statistical significance test used for analyzing the contingency

tables (McCrum-Gardner, 2008). In practice, the method could be used, even sample sizes are small, since, the method is valid for all sample sizes (Fisher, 1992). To select the appropriate data analysis method, it is crucial to define the experimental setting of the conducted experiment and data characteristics (Table 4.18). According to the dataset in the experiment, Fisher Exact test is selected as the most appropriate method (Figure 4.4).

Table 4.18. Experiment and Data Features for Mutually Exclusive Groups.

Experiment	**Hazard Identification Experiment (Mutually Exclusive Groups)**
Data Measure Type of Dependent Variable and Independent Variables	Categorical, Ordinal Data
Data Characteristics	Binary, Discrete Data
Experimental Procedure	Independent Samples
Number of Groups	2
Potential Analysis Method	Fisher Exact Test

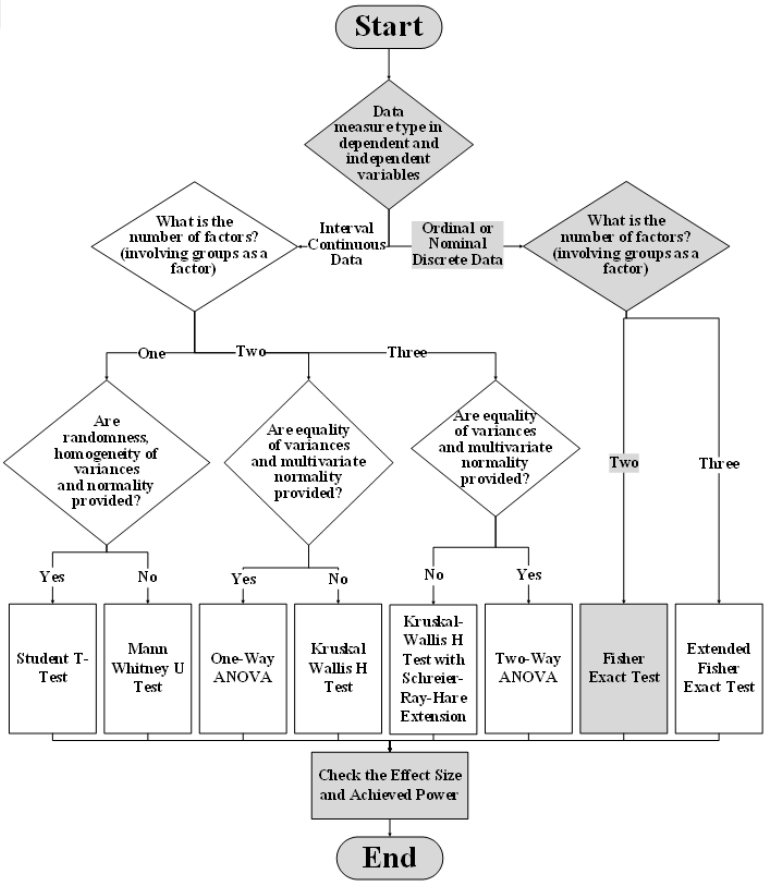


Figure 4.4. The Appropriate Data Analysis Method for Hazard Identification Experiment for Independent Groups.

4.3.1. Descriptive Analyses

According to the previous literature, the major assumptions of chi-square related tests such as McNemar and Fisher Exact test are quite similar to each other. Except for experimental settings such as paired or independent dataset usage, randomness is the only assumption on these tests (McCrum-Gardner, 2008). As the preliminary analysis of data, randomness test and power tests were conducted.

4.3.1.1. Testing the Power. The achieved power is an important parameter that must be taken into consideration in McNemar Test (Kılıç, 2014). To analyze the effect size, power analysis was performed. The analysis results were illustrated in Table 4.19, Table 4.3.1.1, and Table 4.21.

Table 4.19. The Summary of the Power Test for 1st Group.

Risk	**Trial**	**Odds Ratio**	**Sample Size**	**The Proportion Of Discordant Pairs**	**Lower Critical N**	**Upper Critical N**	**Actual α**	**Power (1-β)**
Fall from height	Lecture Notes V-SAFE	40.11	17	0.08	1	2	0.5000	0.9757
Electrocution	Lecture Notes V-SAFE	441.00	17	0.03	1	2	0.5000	0.9977
Building / soil collapse	Lecture Notes V-SAFE	0.00	17	0.58	0	0	0.0156	0.9881
Accident by falling object	Lecture Notes V-SAFE	0.09	17	0.47	0	0	0.0313	0.6499
Heavy equipment accidents	Lecture Notes V-SAFE	4.59	17	0.19	2	3	0.25	0.6742
Fire or explosion	Lecture Notes V-SAFE	133	17	0.06	1	2	0.5	0.9925
Traffic accidents on site	Lecture Notes V-SAFE	1.48	17	0.28	3	4	0.125	0.2125
Materials-based risks	Lecture Notes V-SAFE	441	17	0.03	7	7	0.5	0.9977
Total	Lecture Notes V-SAFE	3.8	17	0.2	2	3	0.25	0.6267

Table 4.20. The Summary of the Power Test for 2nd Group

Risk	**Trial**	**Odds Ratio**	**Sample Size**	**The Proportion of Discordant Pairs**	**Lower Critical N**	**Upper Critical N**	**Actual α**	**Power (1-β)**
Fall from height	Lecture Notes V-SAFE	6.22	17	0.36	10	10	0.0461	0.9063
Electrocution	Lecture Notes V-SAFE	25	17	0.19	7	7	0.0078	0.7599

Table 4.20. The Summary of the Power Test for 2nd Group (cont.).

Risk	Trial	Odds Ratio	Sample Size	The Proportion of Discordant Pairs	Lower Critical N	Upper Critical N	Actual α	Power (1-β)
Building / soil collapse	Lecture Notes V-SAFE	2.03	17	0.42	12	12	0.0384	0.3495
Accident by falling object	Lecture Notes V-SAFE	12.29	17	0.28	9	9	0.0327	0.9555
Heavy equipment accidents	Lecture Notes V-SAFE	145	17	0.11	4	5	0.0625	0.9729
Fire or explosion	Lecture Notes V-SAFE	25	17	0.19	7	7	0.0078	0.7599
Traffic accidents on site	Lecture Notes V-SAFE	4.63	17	0.33	10	10	0.0193	0.6383
Materials-based risks	Lecture Notes V-SAFE	79.55	17	0.17	7	7	0.0078	0.9163
Total	Lecture Notes V-SAFE	10.8	17	0.24	8	8	0.0195	0.8263

Table 4.21. The Summary of the Power Test for the Improvement Between Two Independent Groups.

Risk	Trial	Odds Ratio	Sample Size	The Proportion of Discordant Pairs	Lower Critical N	Upper Critical N	Actual α	Power (1-β)
Fall from height	Group 1 Group 2	9.24	27	0.45	11	11	0.0287	0.9591
Electrocution	Group 1 Group 2	37.15	27	0.24	7	7	0.0078	0.8303
Building / soil collapse	Group 1 Group 2	5.07	27	0.52	12	12	0.0384	0.8909
Accident by falling object	Group 1 Group 2	18.79	27	0.34	9	9	0.0107	0.9123
Heavy equipment accidents	Group 1 Group 2	87	27	0.14	5	5	0.0313	0.9445
Fire or explosion	Group 1 Group 2	11.21	27	0.28	8	8	0.0195	0.8355
Traffic accidents on site	Group 1 Group 2	1.21	27	0.52	12	12	0.0384	0.0822
Materials-based risks	Group 1 Group 2	47.73	27	0.21	7	7	0.0078	0.8649
Total	Group 1 Group 2	18.02	27	0.33	9	9	0.0107	0.9061

The power analysis results show that the majority of hazard categories contain a quite high type 1 error (actual α) in 1st group. Except building/soil collapse and accident by falling object hazard categories, type 1 error rates are higher than 0.05. Therefore, evaluating the hazard sources, except building/soil collapse and accident by a falling object in the 1st group could be trivial and statistically insignificant.

In group 2, the power analysis results show that the power of building and soil collapse hazard category is 0.3495, which is quite less than the objected power; 0.80. Secondly, heavy equipment accidents type 1 error rate is 0.0625, which is slightly higher than 0.05. Although the error rate is higher than 0.05, this hazard category could be analyzed due to the low difference. Thirdly, traffic accidents on-site hazard category is much less than the objected power. Therefore, it is not statistically meaningful to analyze this hazard source. Finally, electrocution and fire or explosion hazard sources are little less than objected power, which is 0.80. Therefore, except building and soil collapse, and traffic accidents on-site hazard types, it is decided to analyze all-hazard sources.

Finally, the power analysis results show that the power and type error rate significantly achieved the expected results between the improvements of two groups. Therefore, according to power analysis results, evaluating the improvement between the two groups are statistically meaningful.

4.3.1.2. Testing the Randomness. To check, whether or not randomness is provided in the data, Wald-Wolfowitz Runs Test is performed. The analysis results are illustrated in Table 4.3.1.2 and Table 4.23.

Table 4.22. Randomness Test Results of Group 1.

Hazard	**Training Method**	**Sigma**	**Z**	**P > |Z|**	**Randomness**
Fall from a height	V-SAFE	0.4	0.5	0.6171	Yes
Fall from a height	Lecture Notes	0.4	0.5	0.6171	Yes
Fall from a height	Improvement	N.A.	N.A.	N.A.	No
Electrocution	V-SAFE	N.A.	N.A.	N.A.	No
Electrocution	Lecture Notes	N.A.	N.A.	N.A.	No
Electrocution	Improvement	N.A.	N.A.	N.A.	No
Building/soil collapse	V-SAFE	N.A.	N.A.	N.A.	No
Building/soil collapse	Lecture Notes	N.A.	N.A.	N.A.	No
Building/soil collapse	Improvement	N.A.	N.A.	N.A.	No
Accident by the falling object	V-SAFE	0.88	-2.49	0.0123	No
Accident by the falling object	Lecture Notes	0.88	-2.49	0.0123	No
Accident by the falling object	Improvement	N.A.	N.A.	N.A.	No
Heavy equipment accidents	V-SAFE	1,22	-0.16	0.8700	Yes
Heavy equipment accidents	Lecture Notes	1.22	-0.16	0.8700	Yes
Heavy equipment accidents	Improvement	N.A.	N.A.	N.A.	No
Fire or explosion	V-SAFE	0.4	0.5	0.6171	Yes
Fire or explosion	Lecture Notes	N.A.	N.A.	N.A.	No
Fire or explosion	Improvement	0.4	0.5	0.6171	Yes

Table 4.22. Randomness Test Results of Group 1 (cont.).

Hazard	Training Method	Sigma	Z	P > |Z|	Randomness
Traffic accident on site	V-SAFE	1.42	-0.56	0.5741	Yes
Traffic accident on site	Lecture Notes	1.22	-0.16	0.8700	Yes
Traffic accident on site	Improvement	1.22	-0.98	0.3261	Yes
Materials-based risks	V-SAFE	N.A.	N.A.	N.A.	No
Materials-based risks	Lecture Notes	N.A.	N.A.	N.A.	No
Materials-based risks	Improvement	N.A.	N.A.	N.A.	No
Total	V-SAFE	0.92	0.65	0.5127	Yes
Total	Lecture Notes	1.42	-0.56	0.5741	Yes
Total	Improvement	1.42	0.14	0.8883	Yes

Table 4.23. Randomness Test Results of Group 2.

Hazard	Training Method	Sigma	Z	P > |Z|	Randomness
Fall from a height	Lecture Notes	1.81	-0.42	0.673	Yes
Fall from a height	V-SAFE	0.32	0.37	0.715	Yes
Fall from a height	Improvement	1.63	-1.26	0.2079	Yes
Electrocution	Lecture Notes	1.63	-0.04	0.9713	Yes
Electrocution	V-SAFE	0.32	0.37	0.715	Yes
Electrocution	Improvement	1.81	-0.42	0.673	Yes
Building/soil collapse	Lecture Notes	1.63	-0.65	0.5172	Yes
Building/soil collapse	V-SAFE	1.99	0.27	0.7901	Yes
Building/soil collapse	Improvement	1.1	-1.76	0.0785	Yes
Accident by the falling object	Lecture Notes	1.99	1.27	0.2034	Yes
Accident by the falling object	V-SAFE	0.32	0.37	0.715	Yes
Accident by the falling object	Improvement	1.99	1.27	0.2034	Yes
Heavy equipment accidents	Lecture Notes	1.1	0.96	0.3372	Yes
Heavy equipment accidents	V-SAFE	N.A.	N.A.	N.A.	No
Heavy equipment accidents	Improvement	1.1	0.96	0.3372	Yes
Fire or explosion	Lecture Notes	1.63	1.19	0.2351	Yes
Fire or explosion	V-SAFE	0.32	0.37	0.715	Yes
Fire or explosion	Improvement	1.81	-0.42	0.673	Yes
Traffic accident on site	Lecture Notes	1.1	0.96	0.3372	Yes
Traffic accident on site	V-SAFE	1.99	1.77	0.0759	Yes
Traffic accident on site	Improvement	1.81	-0.97	0.3301	Yes
Materials-based risks	Lecture Notes	1.63	1.8	0.072	Yes
Materials-based risks	V-SAFE	N.A.	N.A.	N.A.	No
Materials-based risks	Improvement	1.63	1.8	0.072	Yes
Total	Lecture Notes	1.6	0.73	0.4622	Yes
Total	V-SAFE	1.93	-0.12	0.903	Yes
Total	Improvement	1.63	-1.26	0.2079	Yes

Analysis results highlight that the majority of the dataset in the first group have not provided the randomness. Therefore, except traffic accident on site, falling from a height, heavy equipment accidents, fire or explosion, and total categories, comparing the differences in the first group is statistically insignificant. Also, the majority of datasets in group 1 fail to provide necessary type 1 error rate. Therefore, the evaluation of group 1 would not make a statistical meaning. So, it is decided to not analyze the data in the first group for both independent samples and paired tests.

Except for building/soil collapse category, all hazard sources provide adequate type 1 error rate and targeted power in the second group. Besides, all hazard categories in the second group provide randomness. Therefore, it is statistically meaningful to analyze all datasets, except building/soil collapse category in group 2. However, due to the inadequacies in type 1 error rates, power, and randomness in the first group hazard categories, it is insignificant to conduct an independent samples analysis.

4.3.2. Analysis Results

In statistics, the McNemar is a statistical test used in paired binomial data. Due to the dataset of each risk source is binomial, it is determined to use the McNemar test to evaluate whether the row and column marginal frequencies are equal. The analysis results of the McNemar test is illustrated in Table 4.24. Consequently, the status of the hypotheses is represented as following (Table 4.25).

Table 4.24. The summary of McNemar test analysis results for 2nd Group.

Risk	**Trial**	**Mean**	**St. Dev.**	**Sample Size**	**Exact Significance**	**Impact**
Fall from height	Lecture Notes	0.647	0.493	17	0.031	**
	V-SAFE	0.941	0.243	17		
Electrocution	Lecture Notes	0.294	0.47	17	0.001	***
	V-SAFE	0.941	0.243	17		
Building / soil collapse	Lecture Notes	0.294	0.47	17	0.25	N.S.
	V-SAFE	0.471	0.514	17		
Accident by falling object	Lecture Notes	0.471	0.514	17	0.008	***
	V-SAFE	0.941	0.243	17		
Heavy equipment accidents	Lecture Notes	0.176	0.393	17	0	***
	V-SAFE	1	0	17		
Fire or explosion	Lecture Notes	0.294	0.47	17	0	***
	V-SAFE	0.941	0.243	17		
Traffic accidents on site	Lecture Notes	0.176	0.393	17	0.25	N.S.
	V-SAFE	0.529	0.514	17		
Materials-based risks	Lecture Notes	0.294	0.47	17	0	***
	V-SAFE	1	0	17		
Total	Lecture Notes	2.647	1.539	17	0	***
	V-SAFE	6.765	0.903	17		
In the first group, trainees firstly identified hazard sources through V-SAFE. Then, they identified hazard sources through lecture notes.						
***: $\rho < 0.01$. N.S.= not statistically significant.						
***: $\rho < 0.01$. N.S.= not statistically significant.						

Table 4.25. The status of the hypotheses.

Hypotheses	Abbreviation	Status
Trainees trained by V-SAFE identify the risk of falling from height more effectively than trainees trained by the conventional safety training methods.	H3.1.	Accepted ($\rho < 0.05$)
Trainees trained by V-SAFE identify the risk of electrocution more effectively than trainees trained by the conventional safety training methods.	H3.2.	Accepted ($\rho < 0.01$)
Trainees trained by V-SAFE identify the risk of building/soil collapse more effectively than trainees trained by the conventional safety training methods.	H3.3.	Rejected
Trainees trained by V-SAFE identify the risk of accident from the falling object more effectively than trainees trained by the conventional safety training methods.	H3.4.	Accepted ($\rho < 0.01$)
Trainees trained by V-SAFE identify the risk of heavy equipment accidents more effectively than trainees trained by the conventional safety training methods.	H3.5.	Accepted ($\rho < 0.01$)
Trainees trained by V-SAFE identify the risk of fire or explosion accidents more effectively than trainees trained by the conventional safety training methods.	H3.6.	Accepted ($\rho < 0.01$)
Trainees trained by V-SAFE identify the traffic accidents risk more effectively than trainees trained by the conventional safety training methods.	H3.7.	Rejected
Trainees trained by V-SAFE identify the risk of materials-based risks more effectively than trainees trained by the conventional safety training methods.	H3.8.	Accepted ($\rho < 0.01$)
Trainees trained by V-SAFE identify commonly encountered risks more effectively than trainees trained by the conventional safety training methods.	H3.9.	Accepted ($\rho < 0.01$)

4.4. Eye-Movement Behavior Comparison Experiment

This experiment is conducted for evaluating the Statement 4: *Statement 4: Trainees? eye-tracking behavior in V-SAFE are more effective than the trainees on a real construction site.* The analysis results are illustrated in Appendix A. The descriptive statistics of the analysis is shown in Table 4.26.

Table 4.26. Experiment Results of Eye-Movement Behavior Comparison Experiment.

	Time To First Fixation	Total Fixation Duration	Fixation Count	First Fixation Duration
Real Construction Site - Mean	6.28	1.95	8.87	0.15
Real Construction Site - St. Deviation	11.12	3.01	11.18	0.13
Real Construction Site - N	91	91	91	91
V-SAFE - Mean	19.24	3.35	10.65	0.29
V-SAFE - St. Deviation	25.58	3.79	12.51	0.35
V-SAFE - N	127	127	127	127

It was noted that the dataset contains interval and continuous data. Moreover, the experiment was conducted through two different groups. To analyze the differences among mutually exclusive groups, two data analysis methods such as Student t-test and Mann-Whitney u test existed in the literature (see Figure 3.16). To select the appropriate data analysis method, it is important to define the experimental setting of the conducted experiment and data characteristics (Table 4.27).

Table 4.27. Experiment and Data Features.

Experiment	Eye-Movement Behavior Comparison Experiment
Data Measure Type of Dependent Variable and Independent Variables	Interval Data
Data Characteristics	Continuous Data
Experimental Procedure	Independent Samples
Number of Independent Groups	2
Potential Analysis Methods	Student t-test, Mann-Whitney U test

The analysis of student t-test contains various assumptions such as normality, homogeneity, and randomness, while, the only assumption of Mann Whitney-U Test is randomness. Unless these assumptions are met, then the analysis method will not run statistically significant results. The purpose of analysis of the collected data will be decided upon the assessment of the assumptions of the available techniques, and the most appropriate method of analysis will be determined. To select the appropriate data analysis method, several descriptive analyses were conducted, which explained in the following section.

4.4.1. Descriptive Analyses

The assumption of normality always requires to be evaluated for parametric tests, since the validity depends on the normality of the data (Ghasemi and Zahediasl, 2012). Also, the assumption of homogeneity is another critical parameter that needs to be evaluated for parametric tests (Cuzick, 1995).

4.4.1.1. Testing the Normality of the Data. As the initial analysis, it is checked the normality of the eye-tracking data by using the Kolmogorov Smirnov test illustrated in Table 4.28.

Table 4.28. Kolmogorov Smirnov Test Results.

Metric	Experiment	Unit	Count	Mean	St. Dev.	Skewness	Kurtosis	D	ρ - value
Time To First Fixation	Real Construction Site	Second	91	6.28	11.12	3.11101	11.45874	0.29	*0.00001* ***
	V-SAFE	Second	127	19.24	25.58	1.68749	2.860908	0.23	*0.00001* ***
Total Fixation Duration	Real Construction Site	#	91	1.95	3.01	2.89561	11.146205	0.26	*0.00001* ***
	V-SAFE	#	127	3.35	3.79	1.854631	4.029669	0.19	*0.0002* ***
Fixation Count	Real Construction Site	Second	91	8.87	11.18	2.340675	6.519743	0.24	*0.00004* ***
	V-SAFE	Second	127	10.65	12.51	2.294403	6.341663	0.22	*0.00001* ***
First Fixation Duration	Real Construction Site	Second	91	0.15	0.12	1.542132	2.825057	0.13	*0.08468* *
	V-SAFE	Second	127	0.29	0.35	3.895181	18.57625	0.22	*0.00001* ***

4.4.1.2. Testing the Homogeneity of the Data. To check, whether or not homogeneity is provided in the data, the Levene test is performed. The analysis results are illustrated in Table 4.4.1.2 and Table 4.30.

Table 4.29. The summary of data variables.

Eye Tracking Metric		Sample Variances		
		c	**d**	**Total**
	N	91	127	218
Time to First Fixation	ΣX	639.7129	2543.166	3182.879
	Mean	7.0298	20.0249	14.6
	ΣX^2	11135.01	82424.61	93559.62
	Std.Dev.	8.5881	15.8109	14.7308
Total Fixation Duration	N	91	127	218
	ΣX	181.3115	363.3367	544.6482
	Mean	1.9924	2.8609	2.498
	ΣX^2	814.35	1811.13	2625.48
	Std.Dev.	2.2438	2.4747	2.4142
Fixation Count	N	91	127	218
	ΣX	714.1758	1125.102	1839.278
	Mean	7.8481	8.8591	8.437
	ΣX^2	11252.42	19706.76	30959.17
	Std.Dev.	7.9215	8.7919	8.4355

Table 4.29. The summary of data variables (cont.).

Eye Tracking Metric		Sample Variances		
		c	**d**	**Total**
	N	91	127	218
Time to First Fixation First Fixation Duration	N	91	127	218
	ΣX	8.591	25.0219	33.6129
	Mean	0.0944	0.197	0.154
	ΣX^2	1.4626	15.6864	17.149
	Std.Dev.	0.0851	0.2922	0.2348

Table 4.30. The variables descriptive and the results of Levenes' test of Homogeneity.

Eye Tracking Metric	Source	SS	DF	MS	F	ρ - value
Time to First Fixation	Between-treatments	8952.5951	1	8952.5951	50.70716	0.00001
	Within-treatments	38135.849	216	176.5549		
	Total	47088.444	217			
Total Fixation Duration	Between-treatments	39.9864	1	39.9864	7.0521	0.008507
	Within-treatments	1224.7517	216	5.6701		
	Total	1264.7381	217			
Fixation Count	Between-treatments	54.1852	1	54.1852	0.76065	0.76065
	Within-treatments	15386.895	216	71.2356		
	Total	15441.081	217			
First Fixation Duration	Between-treatments	0.5582	1	0.5582	10.56965	0.001334
	Within-treatments	11.408	216	0.0528		
	Total	11.9663	217			

4.4.1.3. Testing the Randomness. To check, whether or not randomness is provided in the data, Wald-Wolfowitz Runs Test is performed. The analysis results are illustrated in Table 4.31.

Table 4.31. The summary of randomness test analysis results.

	V-SAFE							
	Time Fixation Fixation	**Total Fixation Duration**	**Fixation Count**	**First Fixation Duration**	**Time To First Fixation**	**Total Fixation Duration**	**Fixation Count**	**First To First Duration**
Test Value	1.48	0.55	4	0.12	6.43	1.99	6	0.22
Cases < Test Value	45	45	43	45	63	63	60	63
Cases >= Test Value	46	46	48	46	64	64	67	64
Total Cases	91	91	91	91	127	127	127	127
Number of Runs	40	44	48	48	60	62	60	67
Z	-1.369	-0.526	0.346	0.317	-0.801	-0.445	-0.77	0.446
(2-tailed) Asymp. Sig.	0.171	0.599	0.729	0.751	0.423	0.656	0.441	0.655

Based on the Kolmogorov-Smirnov analysis results, all eye tracking metrics' skewness and kurtosis values are not close to 0. In addition, corresponding ρ - value was less than 0.01. Therefore, except for the first fixation duration of the trainees on a real construction site, all eye tracking metrics are not normally distributed. Moreover, except fixation count, all eye tracking metrics do not provide homogeneity. Also, Wald-Wolfowitz runs test results show that the significance values are much higher than 0.05 (Table 4.30). According to Wald-Wolfowitz runs test results, one can conclude that randomness is significantly provided on each metrics in the dataset. Therefore, it is decided to use non-parametric analysis methods. To compare the search patterns of the trainees between V-SAFE and real construction site interaction, it is used the Mann-Whitney U test (Figure 4.5).

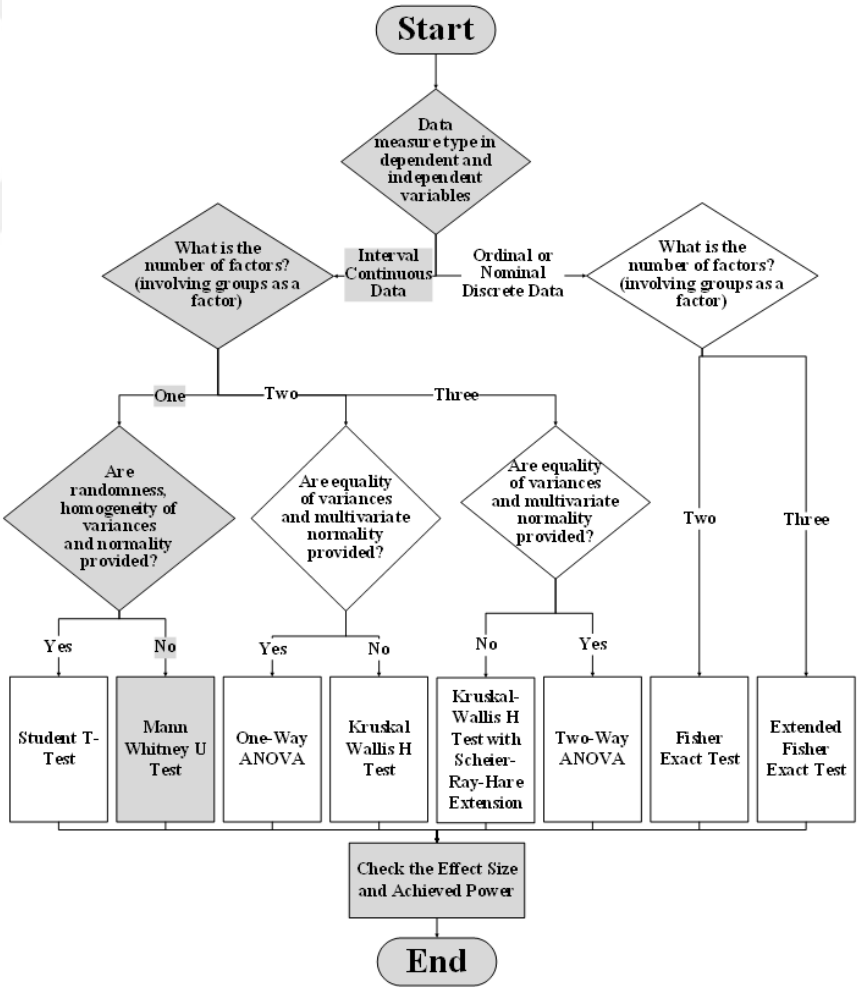


Figure 4.5. The Appropriate Data Analysis Method for Eye-Movement Behavior Comparison Experiment.

4.4.1.4. Testing the Effect Size and Power. As mentioned, the effect size is a concept that must be taken into consideration when calculating the minimum required sample size when planning any study (Kılıç, 2014). Even if the difference of means is viable and statistically significant, the results could be trivial, when the effect size is small. In order to analyze the effect size, Cohen t-test is performed. The analysis results are illustrated in Table 4.32.

Table 4.32. The summary of Cohen t-test analysis results

Eye Tracking Metric	Environment	Mean	St. Dev.	Sample Size	Effect Size	Power
Time to First Fixation	Real Construction Site	6.2765	11.123	91	0.6574	0.9991
	V-SAFE	19.2426	25.576	127		
Total Fixation Duration	Real Construction Site	1.9482	3.008	91	0.4089	0.907
	V-SAFE	3.3475	3.7913	127		
Fixation Count	Real Construction Site	8.8681	11.1815	91	0.1505	0.2903
	V-SAFE	10.6535	12.5061	127		
First Fixation Duration	Real Construction Site	0.1532	0.1275	91	0.4986	0.9758
	V-SAFE	0.2854	0.3528	127		

Cohen analysis results show that the effect size of fixation count is quite small. Moreover, the power of fixation count is 0.2903, which is quite low. Therefore, one can conclude that evaluating the impact of fixation count is not statistically meaningful. Secondly, the first fixation duration, time to first fixation and total fixation duration are quite close to medium effect size, while a high level of power is achieved in these metrics' analysis. Therefore, improving the sample size might be considered as a suitable method to enhance the effect size. However, due to the very high level of achieved power, the impact of effect size could be neglected.

4.4.2. Analysis Results

To compare the difference in the eye-tracking behavior of the trainees, non-parametric independent samples t-test analysis is conducted. Table 4.33 illustrates the detailed independent samples non-parametric t-test results.

Table 4.33. Mann Whitney-U test analysis results.

	Time To First Fixation	**Total Fixation Duration**	**Fixation Count**	**First Fixation Duration**
Real Construction Site - Mean	6.28	1.95	8.87	0.15
Real Construction Site - St. Deviation	11.12	3.01	11.18	0.13
Real Construction Site - N	91	91	91	91
V-SAFE - Mean	19.24	3.35	10.65	0.29
V-SAFE - St. Deviation	25.58	3.79	12.51	0.35
V-SAFE - N	127	127	127	127
Mann-Whitney U	3861.5	3945.5	4678	5031
Z	-4.18	-3.99	-2.4	-1.64
Asymptote Significance (2-tailed)	0	0	0.02	0.1
Exact Significance (2-tailed)	0	0	0.02	0.1
Exact Significance (1-tailed)	0	0	0.01	0.05

Point probability, significance (1-tailed) values on each metrics indicate that there is a statistical difference between the workers' eye-tracking data on the real construction site and V-SAFE training. However, since power value and effect sizes are significantly small in fixation count metric, analysis results could not be considered as significant. Therefore, accepted hypotheses are illustrated in Table 4.34.

Table 4.34. The status of the hypotheses in eye-movement behavior comparison.

Hypotheses	**Abbreviation**	**Status**
The time to first fixation durations of participants using V-SAFE are significantly lower than the time to first fixation durations of the same participants on the real construction site.	H4.1.	Rejected ($\rho < 0{,}01$)
The total fixation durations of participants using V-SAFE are significantly higher than the total fixation durations of the same participants on the real construction site.	H4.2.	Accepted ($\rho < 0{,}01$)
The first fixation durations of participants using V-SAFE are significantly higher than the first fixation durations of the same participants on the real construction site.	H4.3.	Accepted ($\rho < 0{,}01$)
The fixation count of participants using V-SAFE are significantly higher than the fixation count of the same participants on the real construction site.	H4.4.	Rejected due to small effect size and power

4.5. Construction Site and Gaming Experience Experiment

This experiment is conducted for evaluating the following hypotheses: *Hypothesis 5.1: Both trainees with and without site experience could use virtual construction safety training tool in a similar and effective manner Hypothesis 5.2: Both trainees with and without computer gaming experience could use virtual construction safety training tool in a similar and effective manner.*

Table 4.35 illustrates the experimental results, including the safety training performance together with the construction site and gaming experience. Table 4.36 summarizes descriptive results, including the safety training performance, ranks along with the gaming and construction site experience.

Table 4.35. Analysis results of the trainees based on construction site experience.

Trainee	Site Experience	Gaming Experience	Safety Training Performance
1	No	1	1.00
2	No	1	0.96
3	No	1	0.95
4	No	2	0.8
5	No	2	1.00
6	No	2	1.00
7	No	3	1.00
8	No	3	1.00
9	No	3	0.50
10	No	4	0.92
11	No	4	1.00
12	No	4	0.69
13	Yes	1	0.96
14	Yes	1	0.92
15	Yes	1	0.78
16	Yes	2	1.00
17	Yes	2	0.61
18	Yes	2	0.92
19	Yes	3	0.93
20	Yes	3	0.90
21	Yes	3	1.00
22	Yes	4	1.00
23	Yes	4	1.00
24	Yes	4	1.00

Table 4.36. Descriptive results.

Factor	Level	N	Mean	Std. Dev.
Construction Site Experience Construction Gaming Experience	No 1	3	0.9700	0.02646
Construction Site Experience Construction Gaming Experience	No 2	3	0.9333	0.11547
Construction Site Experience Construction Gaming Experience	No 3	3	0.8333	0.28868
Construction Site Experience Construction Gaming Experience	No 4	3	0.8700	0.16093
Construction Site Experience Construction Gaming Experience	No Total	12	0.9017	0.15971
Construction Site Experience Construction Gaming Experience	Yes 1	3	0.8867	0.09452
Construction Site Experience Construction Gaming Experience	Yes 2	3	0.8433	0.20599
Construction Site Experience Construction Gaming Experience	Yes 3	3	0.9433	0.05132
Construction Site Experience Construction Gaming Experience	Yes 4	3	1.0000	0.00000
Construction Site Experience Construction Gaming Experience	Yes Total	12	0.9183	0.11668
Construction Site Experience Construction Gaming Experience	Total 1	6	0.9283	0.07705
Construction Site Experience Construction Gaming Experience	Total 2	6	0.8883	0.15728
Construction Site Experience Construction Gaming Experience	Total 3	6	0.8883	0.19498
Construction Site Experience Construction Gaming Experience	Total 4	6	0.9350	0.12422
Construction Site Experience Construction Gaming Experience	Total Total	24	0.9100	0.13705

From the first overview, the average of trainees' safety performance in V-SAFE is quite similar to each other. The average of trainees' score without construction site experience was 0.9017, while the average of trainees with site experience was 0.9183. The difference between the trainees with and without site experience was 0.0166.

It was noted that the dataset contains interval and continuous data. Moreover, there are two different factors evaluated on the dataset, which could contain potential correlations with each other. Comparing safety performance through one factor could

result in neglecting the potential correlations between factors. Therefore analyzing the impact of two factors to the dependent variable and evaluating the relation between factors could be an appropriate approach. To analyze the impact of factors, Two-Way ANOVA and Kruskal Wallis H test with Scheirer-Ray-Hare test are suitable methods (see Figure 3.16). To select the appropriate data analysis method, it is essential to define the experimental setting of the conducted experiment and data characteristics (Table 4.37).

Table 4.37. Experiment and Data Features.

Experiment	Construction Site and Gaming Experience Experiment
Data Measure Type of Variables	Dependent Variable: Interval and Continuous Data
Experimental Procedure	Independent Samples
Number of Factors	2
Potential Analysis Methods	Two-Way ANOVA, Kruskal-Wallis H test with Scheirer-Ray-Hare Extension

4.5.1. Descriptive Analyses

As essential requirements of the two-way ANOVA; normality, equality of error variances, effect size are evaluated.

4.5.1.1. Testing the Normality of the Data. To check the data, whether or not the normality is provided, Shapiro-Wilk test for different factors is performed. According to the normality tests, it could be seen three of four variables do not provide normality (Table 4.38).

Table 4.38. Test of normality results.

Shapiro-Wilk				
Factor	**Level**	**Statistic**	**Degrees of Freedom**	**Significance**
Construction Site Experience Construction Gaming Experience	No 1	0.893	3	0.363
Construction Site Experience Construction Gaming Experience	No 2	0.750	3	0.000
Construction Site Experience	No			
Construction Gaming Experience Construction Site Experience	3 No	0.750	3	0.000
Construction Gaming Experience	4	0.928	3	0.480
Construction Site Experience Construction Gaming Experience	Yes 1	0.907	3	0.407
Construction Site Experience Construction Gaming Experience	Yes 2	0.896	3	0.373
Construction Site Experience Construction Gaming Experience	Yes 3	0.949	3	0.567
Construction Site Experience Construction Gaming Experience	Yes 4	-	3	-

4.5.1.2. Testing the Equality of Error Variances. To check, whether or not homogeneity is provided in the data, the Levene test is performed. The analysis results are illustrated in Table 4.39.

Table 4.39. Levene Test of equality of error variances.

Factor	**Levene Statistic**	**df1**	**df2**	**Significance**
Based on Mean	5.749	7	16	0.002
Based on Median	0.556	7	16	0.78
Based on Median and with adjusted degrees of freedom	0.556	7	5.112	0.768
Based on Trimmed Mean	40.825	7	16	0.004

From Table 4.40, it could be noticed that significance (ρ - value) equals to 0.002 based on means, while, it is 0.798 in the median. Previous literature (Levene, 1961) states that if the achieved ρ-value is less than 0.05, the observed differences in sample variances are not likely to occur based on random sampling from a population with

equal variances. Alternatively, the test of equality of error variances could be evaluated through the median, rather than mean. The optimal selection depends on the distribution of dataset (Derrick *et al.*, 2018). In other words, if the distribution is not normally distributed, using the median is optimal (Derrick *et al.*, 2018). According to the normality and equality of error variances tests, Kruskal-Wallis H test with Scheier-Ray-Hare extension should be performed (Figure 4.6).

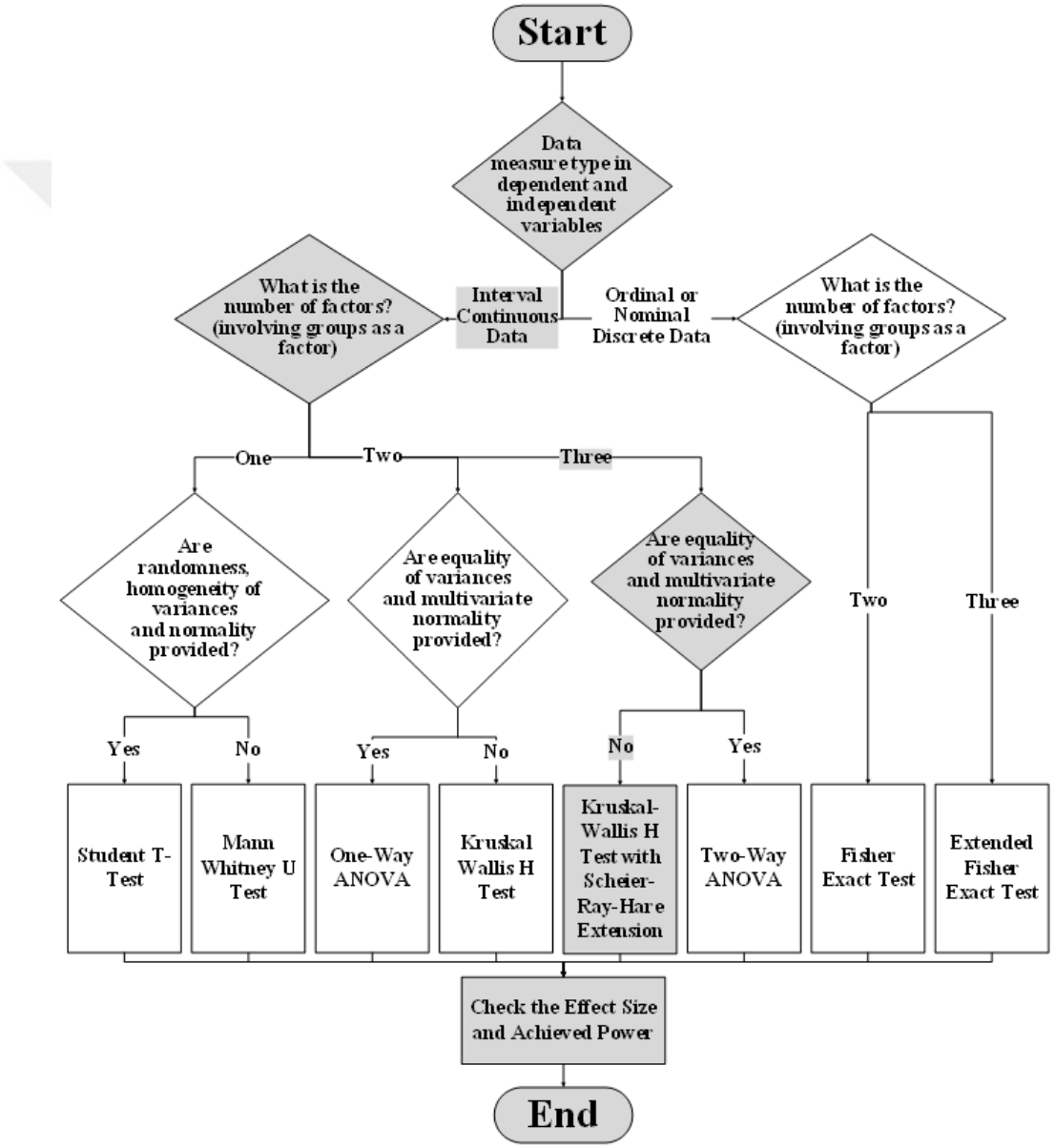


Figure 4.6. The Appropriate Data Analysis Method for Construction Site and Gaming Experiment.

4.5.2. Analysis Results

4.5.2.1. Kruskal-Wallis H Test with Scheier-Ray-Hare Extension. Table 4.40 summarizes the descriptive results, including the safety training performance, ranks together with the gaming and construction site experience. Table 4.41 shows the analysis results of between-subjects effects.

Table 4.40. Kruskal-Wallis Ranks.

Factor	**Level**	**N**	**Mean**	**Std. Dev.**	**Mean Rank**
Construction Site Experience Construction Gaming Experience	No 1	3	0.9700	0.02646	14.1667
Construction Site Experience Construction Gaming Experience	No 2	3	0.9333	0.11547	14.3333
Construction Site Experience Construction Gaming Experience	No 3	3	0.8333	0.28868	13.0000.
Construction Site Experience Construction Gaming Experience	No 4	3	0.8700	0.16093	9.6667
Construction Site Experience Construction Gaming Experience	Yes 1	3	0.8867	0.09452	8.3333
Construction Site Experience Construction Gaming Experience	Yes 2	3	0.8433	0.20599	9.8333
Construction Site Experience Construction Gaming Experience	Yes 3	3	0.9433	0.05132	11.6667
Construction Site Experience Construction Gaming Experience Total	Yes 4	3 24	1.0000	0.00000	19.0000

Table 4.41. Kruskal-Wallis Test Results.

	Sum of Squares	**Degrees of Freedom**	**H**	**ρ-value**	**Significance**
Site Experience	2.0416667	1	0.0451957	0.8316446	-
Gaming Experience	30.75	3	0.6807026	0.8777332	-
Site*Gaming Experience	212.70833	3	4.7086542	0.194417	-
Within	793.5	16			
Total	1039	23			

The results of the Kruskal-Wallis H test show that there is not any relationship between safety training performance, gaming experience, and site experience.

4.5.2.2. Testing the Effect Size. Another critical analysis, which is applied in the ANOVA is effect size. Accordingly, the test of between-subjects is performed to analyze the effect size. The analysis results are illustrated in Table 4.42.

Table 4.42. Achieved Effect Size.

	Partial Eta Square
Safety Training Performance * Experience	0.23628

Based on epsilon squared effect size analysis, it could be seen that the achieved effect size is 0.236284889, which corresponds to a medium effect size (see Table 3.2).

According to the Kruskal-Wallis H test analysis results, the similarity of the training performance for the trainees with and without construction site experience is statistically significant (H = 4.708654, $\rho = 0.194417$). Therefore the status of hypotheses is illustrated as follows (Table 4.43).

Table 4.43. The status of the hypotheses in the construction site and gaming experience experiment.

Both trainees with and without site experience could use virtual construction safety training tool in a similar and effective manner	H5.1.	Accepted
Both trainees with and without computer gaming experience could use virtual construction safety training tool in a similar and effective manner.	H5.2.	Accepted

5. DISCUSSION

5.1. Introduction

As mentioned in chapter 1, the research questions are fulfilled in the thesis concerned whether V-SAFE could provide effective process about safety training parameters defined as safety information transfer, instructional feedback, knowledge development, hazard identification, and hands-on practice. Unlike any other studies, a comparative analysis of the effectiveness of occupational safety training between conventional methods and V-SAFE was performed, and critical factors to provide adequate safety training are discussed. However, significant findings contribute to the understanding of critical factors to provide an effective training method, which is discussed in this chapter.

5.2. Research Motivation

The number of occupational accidents and fatalities have been quite high in the construction industry all around the world. In Turkey, according to official records, 1633 fatal accidents and 587 deaths were occurred in the construction industry in 2017 (SSI, 2019a). Consequently, the construction sector ranks first in terms of the number of occupational accidents in Turkey (SSI, 2019b). Similarly, 3876 fatal occupational accidents occurred in 28 EU countries in 2015 (Eurostat, 2018a), while, around 20% of these fatal accidents took place in the construction industry (Eurostat, 2018b). Moreover, the construction industry had the highest fatal accident rate in the 28 EU countries (Eurostat, 2018b). Just as EU countries and Turkey, 991 fatal accidents were recorded in the American construction industry in 2016 (BLS, 2017). Compared to other sectors, the construction industry had the highest number of occupational deaths in 2016 (BLS, 2017).

Many studies have emphasized the importance of human behavior-oriented factors to provide a safe construction working environment (e.g., Abdelhamid and Everett,

2000, Albert *et al.*, 2014; Carter and Smith, 2006; Shapira and Lyachin, 2009; Lingard and Rowlinson, 2005; Han *et al.*, 2010). Also, several studies (e.g., Cohen et a, 1998; Demirkesen and Arditi, 2015; Wong *et al.*, 1999; Tam and Fung, 2011) highlight the importance of safety training to ensure a safe working environment. A comprehensive literature review (e.g. Burke *et al.*, 2006; Albert *et al.*, 2014; Alliger and Janak, 1989; Frese and Zapf, 1994; Shapira and Elbaz, 2014; Bahn and Barratt-Pugh, 2014) emphasizes that an effective construction safety training method should provide information transfer, knowledge development, hands-on training, behavioral modeling, programmed instructional feedback and hazard identification to the trainees. However, construction companies have used conventional training methods that merely focus on information transfer.

The major objective of this thesis is to develop a conceptual model for evaluating the effectiveness of the developed virtual construction safety training tool, namely V-SAFE, based on effectiveness parameters existed in the literature Figure 5.1. In this context, information transfer, instructional feedback, knowledge development, hazard identification, hands-on practice, behavioral modeling are defined as the parameters of effective safety training. The developed virtual environment contained significant features such as multiplayer support, detailed feedback, tower crane scenarios, realistic construction site simulation, which offer high potential to provide an adequate safety training method. To measure the effectiveness of the developed software, all parameters were validated by five different experimental designs. Analysis results show that all developed research hypotheses were significantly accepted. In the following sections, the relationship between these hypotheses and related statements were discussed.

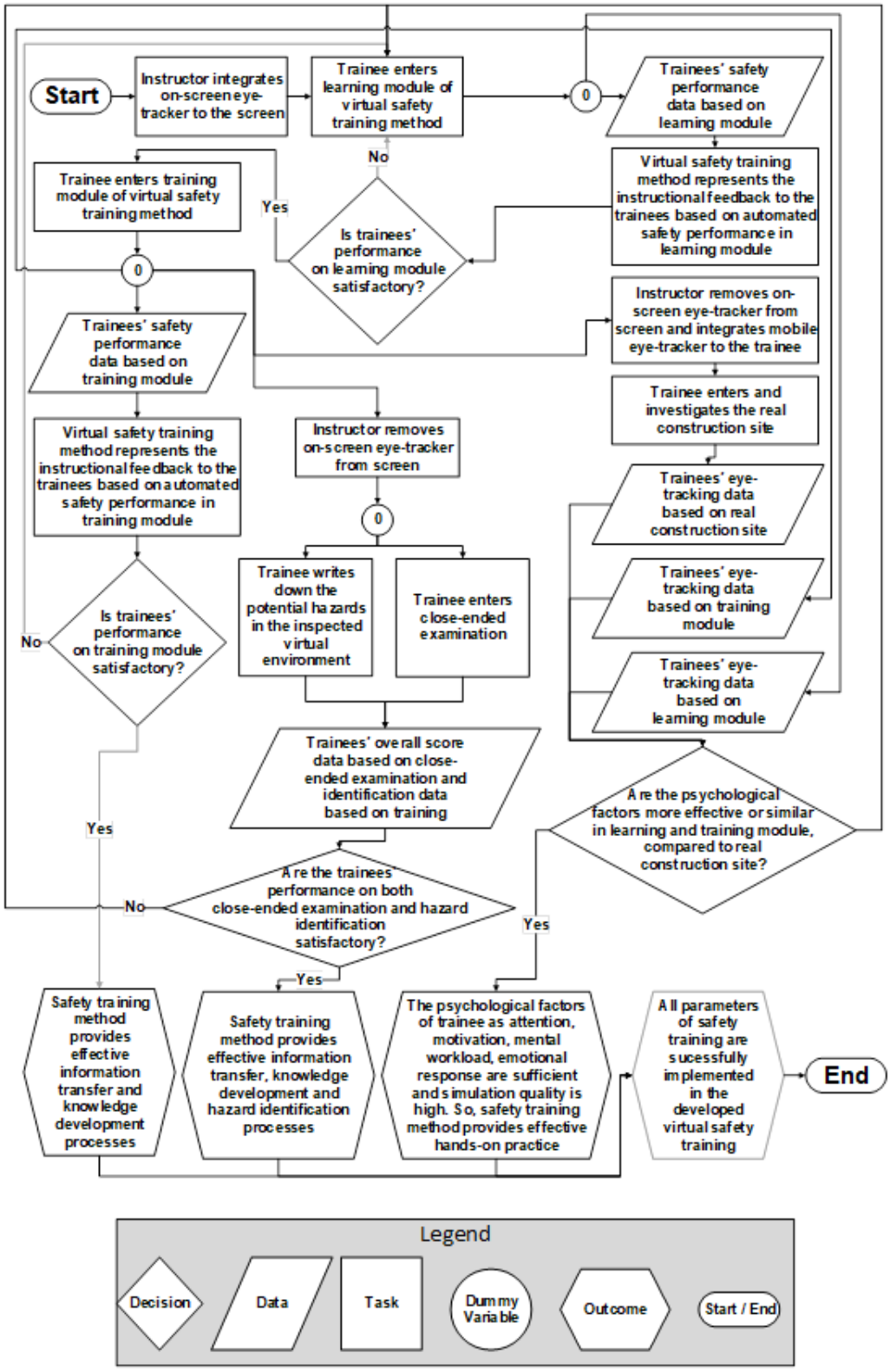


Figure 5.1. Conceptual Model to Evaluate the Effectiveness of Virtual Construction Safety Training Methods.

5.3. Safety Information Transfer

As a primary requirement of the safety training effectiveness, delivering the safety information to the trainees in an appropriate method is compulsory. To provide adequate information representation, Arthur Jr. and colleagues (2003) state that assessing the objectives of training, thus, identifying the essential skills and tasks to be trained is crucial. Based on specific tasks or skills, it is quite likely that any safety training method could be more effective compared to any other (Arthur Jr. *et al.*, 2003). In other words, Arthur Jr. and colleagues (2003) argue that all safety training methods can communicate a specific skill, attitudinal or task information to the trainees; therefore, various training methods could be superior to each other based on the task information. Consistently, Wexley and Latham (2002) state that the determination of task characteristics is quite crucial in finding the most effective safety training method. In this sense, one can conclude that the statement by Wexley and Latham (2002) highlight the importance of information quality context defined by Huang and colleagues (1999). The quality of information could be evaluated by determining what extent the information fits the needs of the consumers (Huang *et al.*, 1999). In order to provide high-quality information in the safety training process, the information must be interpretable, consistent, and relevant to the trainees' needs. In addition, the training method should enhance the mnemonics of trainees (Albert *et al.*, 2014). However, the previous literature (e.g. Frese and Zapf, 1994; Burke *et al.*, 2006) criticizes the conventional safety training (least engaging) methods such as lectures, videos, textbooks fail to provide an adequate safety information transfer process, compared to the highly engaging methods.

To bridge this gap, it is aimed to provide an adequate safety information transfer process through V-SAFE by integrating the learning-by-doing approach. The objective of learning-by-doing strategy is (1) to build a virtual environment design for trainees to experience daily tasks during tower crane tasks in simulation-based scenarios, (2) to establish a base for the trainees to evaluate what safety precautions should be taken on a site, (3) to provide a high quality, understandable, relevant safety information to the trainees. To ensure these parameters, V-SAFE effectiveness is analyzed through

paper-based examination. All questions are directly related to safety information about tasks and the necessary safety information provided by V-SAFE.

The information comprehensibility, which is a two-stage experiment, was conducted to evaluate safety information transfer quality. As the initial stage, all trainees watch video-recorded PowerPoint-based lectures, which could be considered as a quite similar approach for least engaging safety training methods. After completion of conventional safety training, all trainees answer the questions on the close-ended examination. As the second stage, the same trainees entered into V-SAFE and asked to revise their answers in the examination. Analysis results indicate that, when V-SAFE was used after conventional methods, the trainees' mnemonics level are significantly improved (H1.1: $\rho = 0{,}008$, H1.2: $\rho = 0{,}031$).

Several studies (e.g., Li *et al.*, 2012a; Park and Kim, 2013; Le *et al.*, 2014) followed a very similar approach with the experiment design and developed virtual environment. For instance, Li and colleagues (2012a) developed a virtual environment-based safety training method and integrated a close-ended examination procedure. In other words, Li and colleagues (2012a) used the inspection-based environment to establish a question and answer procedure by the integration of questions with the relevant risk source involved in the virtual media. In this procedure, when a trainee moves closer to construction equipment, the specific safety question about the equipment appears, and trainees could answer the question by clicking the correct response in the virtual media. From 20 questions in the database, trainees could answer embedded image-based, text-based, or video-based questions to the virtual objects. Similarly, Park and Kim (2013) developed a virtual environment-based safety training method that contains question and answer game for regular safety education. Workers could decide what-to-do during the risk situation by selecting an option from multiple choice safety measures. Park and Kim (2013) state that utilized virtual-environment-based safety training software to improve the cognition level of trainees by enabling the monitoring of hazardous site elements. Despite the advantages of both approaches, a significant limitation occurs in these studies. A potential weakness of both these methods could be considered as lack of multiplayer support in the developed environments. Without

any multiplayer gaming support, trainees fail to learn safety needs that could be interpreted through collaboration and interaction. Training provided by V-SAFE enables multiplayer support, which allows trainees to gather safety information through a collaborative environment. For instance, the responsible role to assign the load points is bricklayer role in the simulation, while, pointer role could easily learn this information from other roles by interacting with them.

During the evaluation of answers, it was noticed that students were not able to memorize the information definitively after the completion of the PowerPoint presentation. In general, presentations are based on the illustration of graphical elements with verbal explanation (Bartsch and Cobern, 2003), which help the trainees to recall information (ChanLin, 1998; Szabo and Hastings, 2000). However, the graphical elements used in presentations are 2D materials, which are inadequate for translating necessary safety information into a mental picture (Hadikusumo and Rowlinson, 2002). For instance, the lifting process of the load in V-SAFE was based on four-point hinge linking system, which was indicated in the presentation of trainees assigned to pointer and bricklayer roles (totally twelve trainees) as written texts, verbal explanation, and photo (Figure 5.2).



Figure 5.2. The illustration of four-hinge link system in Power Point presentation.

Although the information was provided to twelve trainees, who were assigned to pointer and bricklayer roles in oral, written, and graphical representation together, only five out of these twelve trainees were able to answers the question about four-hinge link system correctly after PowerPoint presentation. During the V-SAFE training, it was observed that three more trainees, who misunderstood during the PowerPoint presen-

tation, actively understand the lifting method by visualizing the lifting process in a 3D virtual environment. Therefore, one can conclude that the visualization of lifting process provides better information quality for the trainees to see the lifting from different angles. Besides, the simulation improves the interaction duration within the object; thus, trainees were able to learn the lifting process, and they were able to improve their spatial mnemonics. Moreover, after the completion of V-SAFE training, trainees were informed by the instructional feedback. As a result, three trainees successfully revised their answers correctly.

In addition, because lectures are instructor-based methods, the duration allocated to transfer some information could be insufficient to memorize the information. For instance, the information required for crane operators to select the right rope type was given in a total of nine seconds, and this time was insufficient for the students to memorize and retain the information correctly. In the experiment, the improper rope selection of the trainee is detected by the virtual tool, and proper feedback is provided to the trainees; thus, trainees could translate this information to their memory without having any time limitation. As a result, the trainees have sufficient amount of time to memorize the information. Acquired information to memory could be considered as between short and long term memory. Short-term memory is a memory that has less than thirty seconds of retention of information (Smith *et al.*, 2013). A very limited amount of information received through detection is retained in short-term memory (Smith *et al.*, 2013). After getting specific feedback from the environment, two trainees out of six, who assigned to crane operator role were able to memorize the information and revised their answers. From these two examples, one could be concluded that V-SAFE improves the memorization level of the trainees by (i) providing better representational information by integrating third dimension, (ii) increasing the interaction duration between information and trainee, (iii) providing specific information to the trainees by instructional feedback. Taken together, V-SAFE is significantly effective (ρ=0.031) in improving the memorization process by meeting the deficits of PowerPoint presentation, and the following hypothesis is accepted: Hypothesis 1.2: *V-SAFE significantly improves the spatial mnemonics of trainees when used after PowerPoint presentations.*

As another essential parameter of the adequate information transfer process, the provided information must be understandable in terms of representational information quality (Al-Hakim, 2007). The PowerPoint lectures are instructor-centered methods, which does not allow the participation of trainees in the process of learning (Daluba, 2013). In this sense, although the PowerPoint-based lectures might be appropriate for the trainees to learn necessary safety information, redundant information might lead to misunderstandings. To bridge this gap, the previous literature (Shakhashiri, 1984) suggests that lectures should be supported by adequate demonstration method, which involves (i) demonstration by the instructor, (ii) allow trainees to practice, what is demonstrated by the instructor. In the experimental setting, students were not able to practice the instructions demonstrated by the instructor before they entered to first examination. During the evaluation of trainees' first examination results, it was noticed that lack of demonstration led to severe misunderstandings. For instance, eight out of twelve trainees misunderstood about the responsibility of assigning loading points. The assignment of loading points is one of the primary duties of the bricklayer role, while, the pointer roles' responsibility is to check the bricklayer, whether the points are correctly assigned. However, eight of twelve trainees in pointer and bricklayer role misunderstood this information, and they thaught crane operator is responsible for assigning loading points. During V-SAFE simulation, trainees completed their tasks including assignment and checked loading points. Therefore, they clearly understood their responsibility, and three of these eight trainees corrected their answers. Quite similarly, trainees assigned to the crane operator role were visually and verbally informed about the load weight, and dimensions should be reported by the trainees of the bricklayer role (Figure 5.2).

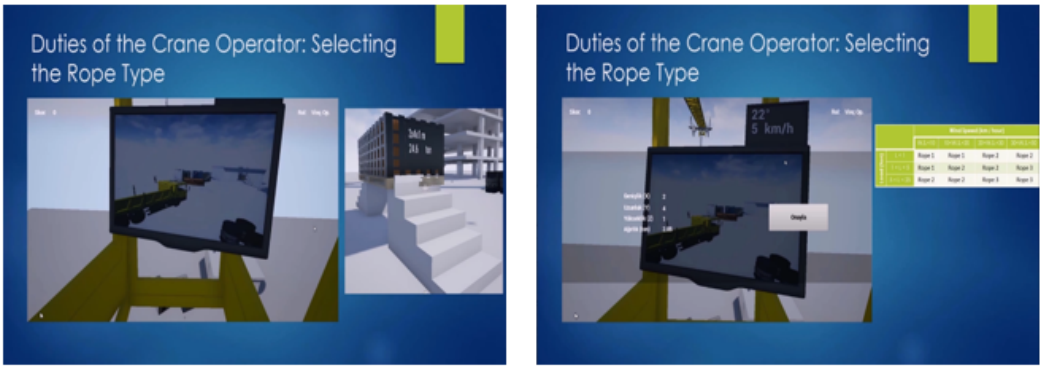


Figure 5.3. The illustration of weight load and dimensions in Power Point presentation.

Before the completion of V-SAFE training, the half of trainees correctly answered the question about informing the load weight and dimensions to the crane operator. After the end of V-SAFE training, two of three trainees, who were not able to answer the question correctly in the first examination, corrected their answers. From these examples, it could be concluded that practicing safety information via V-SAFE significantly reduces misunderstandings (H1.1: $\rho = 0{,}008$).

In brief, the analysis results indicate that V-SAFE has a significant positive impact on safety information transfer quality. Quite consistent with the findings of the previous research, the visual interface of V-SAFE enhances (i) the understandability of trainees by enabling them to monitor the model construction site (Park and Kim, 2013), (ii) improving mnemonics by enhanced information quality and effective instructional feedback (Albert *et al.*, 2014) and (iii) improving the collaboration by providing a multiplayer experiment. Consequently, the following statement is accepted: Statement 1: *V-SAFE significantly improves the quality of safety information when used after conventional safety training methods.*

5.4. Instructional Feedback Mechanism and Knowledge Development

Other parameters of effective safety training are instructional feedback and knowledge development. As mentioned in the previous section, the instructional feedback

allowed learners to correct their mistakes in information comprehensibility experiment. In the repeatable virtual experiment, it is aimed to evaluate the impact of instructional feedback to the knowledge development process.

The previous literature (e.g., Burke *et al.*, 2006) highlights that instructional feedback should be represented to each trainee based on a standardized manner such as on a personal computer or workbook format. In addition, the feedback must be specific for the trainees' weaknesses and safety knowledge level (Burke *et al.*, 2006). This statement is related to the appropriate amount of information parameters defined by Wang and Strong (1996) since the amount of information should be specific for each trainees' weaknesses. The previous literature also (e.g., Kirkpatrick, 1979) highlights that feedback intervention leads learners to change (i) reactions, (ii) learning of skills and knowledge level and (iii) behavior. Therefore, the evaluation of instructional feedback mechanism must be analyzed through these parameters.

To evaluate the changes in trainees' reactions and behavior, a four-step experiment procedure was followed. In the first stage, trainees watched a video-recorded PowerPoint tutorial to ensure that all subject know their roles' responsibilities and essentials of V-SAFE. In the following steps, all participants received V-SAFE training three times. At the end of each stage, trainees receive specific feedback based on automated performance analysis structure. In the previous section, it was noted that instructional feedback was quite assistive for trainees to improve understandability and mnemonics in information transfer (H1.1: $\rho = 0{,}008$, H1.2: $\rho = 0{,}031$). The evaluation of knowledge development should be analyzed through the reactions and the safety behavior of trainees, rather than safety information. V-SAFE uses an automated performance analysis method. The automated performance analysis method of V-SAFE involves three major stages as follows; i) analyzing the safety decisions of the trainees, ii) finding the trainees' safety errors, iii) representing the trainees' safety errors and corrections in a listed format. As a result, V-SAFE provides specific and individual feedback for each trainee based on their decisions and safety behavior during virtual training. In this sense, V-SAFE not only offers effective and need-based feedback but also establishes a base for the trainees to improve safety knowledge and safety behavior

by representing their mistakes and correct them.

To compare the reactions and safety behavior, the differences in trainees' safety training performance were statistically analyzed in three repetitive runs. Between the first and second ($\rho = 0.013$), first and third ($\rho = 0.003$) runs, trainees safety training performance was significantly improved, while there was not any significant improvement between second and third runs ($\rho = 0.263$). The analysis results indicate that two runs repeatable practice is sufficient in V-SAFE. Besides, these results clearly indicate that V-SAFE provides effective instructional feedback for the trainees to change their safety behavior and reactions in the virtual environment. Not only in the lights of these results, but also from the results in the information comprehensibility experiment, one could be concluded that V-SAFE provides an effective training method, in terms of delivering instructional feedback. Therefore, following hypothesis accepted: Hypothesis 2.1: *Trainees using V-SAFE significantly correct their mistakes by repeatable practice.*

Another critical parameter for effective safety training is knowledge development (Burke *et al.*, 2006). As mentioned in section 3.1.2, explicit knowledge is developed through information-based techniques such as pamphlets, lectures and safety notes. Although these methods are able to communicate explicit knowledge effectively, they failed to establish a base for trainees to develop adequate necessary implicit knowledge, since they do not involve any practice. In order to bridge this gap, V-SAFE contains training on the virtual construction site to communicate necessary implicit knowledge. During the investigation of trainees' reaction, it was noticed that trainees were able to improve their behavior in the virtual environment after getting feedback. For instance, putting on safety belt is one of the major precautions to be taken, when workers take place in activities at height. In V-SAFE training, three out of twelve trainees did not put on their safety belt in the first run. However, after getting feedback from V-SAFE, all of them put on their safety belt. As another example, turning on the audio warning system and checking, whether the audio system works, is one of the primary tasks of the crane operator. In V-SAFE training, it was seen that three out of six trainees did not turn on the audio warning system in the first run. In the second run, all trainees corrected their mistakes by activating audio warning system. From these examples,

one can conclude that V-SAFE established a suitable method to improve the safety behavior of trainees by orienting them to correct actions in the virtual environment. In the lights of this research, following hypothesis and statement are accepted:

- Hypothesis 2.2: *V-SAFE significantly improves the implicit knowledge level of trainees by providing repeatable hands-on practice.*
- Statement 2: *V-SAFE provides effective feedback to the trainees, thus, they develop adequate implicit knowledge.*

5.5. Hazard Identification

The effective hazard identification process is quite essential in the construction industry (Carter and Smith, 2006; Huang and Hinze, 2003; Patrucco *et al.*, 2010). The previous literature (Carter and Smith, 2006; Huang and Hinze, 2003; Patrucco *et al.*, 2010) states that the workers' hazard identification level is quite less than targeted level, which leads to the severe on-site accidents. In order to take correct action and manage hazards; identifying the risks properly is a crucial stage. So, the effectiveness of hazard identification stage is quite essential to prevent on-site accidents before they occur.

In order to provide an effective hazard identification process, the previous literature (Chen *et al.*, 2013; Albert *et al.*, 2014) stated that hazard identification is executed through mentally constructing and visualizing the on-site construction tasks, thus, evaluating potential accidents occurred in these activities. The conventional hazard identification methods such as investigating the contract drawings, design drawings, safety notes, etc. cause difficulties to translate information into a mental picture (Hadikusumo and Rowlinson, 2002). Moreover, 2D drawings and safety notes do not contain construction components and processes, where hazards are inherited into them (Hadikusumo and Rowlinson, 2002). As a result, hazard identification could not be undertaken through such methods (Hadikusumo and Rowlinson, 2002). Hadikusumo and Rowlinson (2002) state the necessity of visualizing construction equipment and construction processes to improve hazard identification level. By visualizing the con-

struction processes, and equipment, one can define risks more effectively in terms of integration of the third dimension (Hadikusumo and Rowlinson, 2002).

To evaluate the benefits of 3D environment, a two-stage hazard identification experiment was conducted, and trainees were divided into two groups. The trainees in the first group first identified the risks through V-SAFE and then with conventional occupational safety training methods. The trainees in the second group identified risks firstly with conventional occupational safety training tools and then with V-SAFE. In the first group, around 7% improvement was noticed, while, major assumptions of McNemar test such as the randomness, power, probability of type 1 error, effect size were not in the targeted level in the dataset. Therefore, it was not statistically meaningful to analyze the improvement in the first group. In the second group, there was a significant improvement around 60.8%, while, all assumptions of the McNemar test were met. The analysis results clearly indicate that using V-SAFE significantly improves the hazard identification level of trainees in the accidents: (i) falling from height (H3.1; $\rho = 0.031$), (ii) electrocution (H3.2; $\rho = 0.001$), (iv) accidents by falling object (H3.4; $\rho = 0.008$), (v) heavy equipment accidents (H3.5; $\rho = 0.000$), (vi) fire or explosions (H3.6; $\rho = 0.000$), (viii) materials-based risks (H3.8; $\rho = 0.000$) and total (H3.9; $\rho = 0.000$). The findings make an important contribution to prevent previously occurred accidents. Except soil and structure collapse and traffic accidents on construction site, all-hazard sources frequently encountered in the construction industry were successfully identified by trainees. In addition, the findings offer some important insights into integration of triggers to identify major risk sources by using visualized elements. For instance, several risk sources such as electrocution, fire or explosions risks were mentally visualized by trainees using warning signals occurred in the virtual environment (Figure 5.3). Thus, one can conclude that workers could be able to take correct action by identifying hazards using V-SAFE.

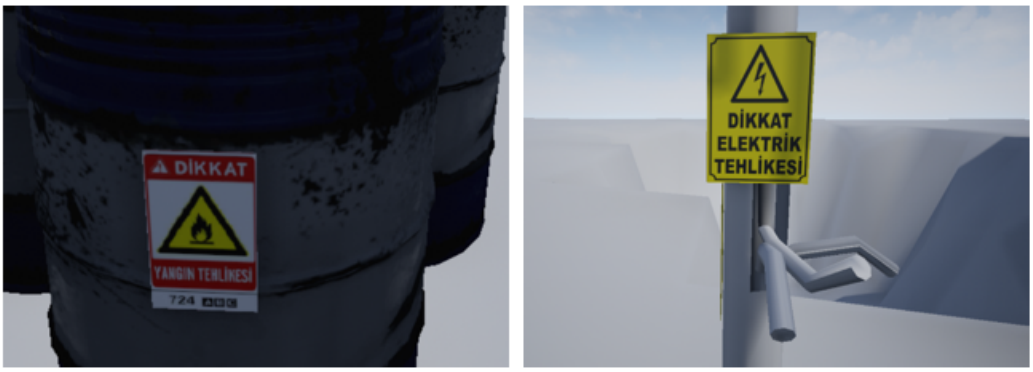


Figure 5.4. The trigger effect of risk identification through visualizing warnings.

However, V-SAFE was found to be ineffective in identifying the risks of soil and structure collapse (H3.3; $\rho = 0.250$). The main reason could be considered as the inadequacy of the visual element in which the soil and structure collapse is defined in the virtual environment (Fig. 3.9). It is thought that the pit defined for identifying the risk of soil collapse in the virtual environment will be effective, but this visual element is not enough to determine this risk. The risk of building collapse was thought to be determined by the collapse of the building model itself, but this visual element was again insufficient. In later versions of the software, a retaining wall to be placed in the ground and damaged columns to be identified in the structure, etc. visuals will be aimed to identify these risks. Another frequently encountered risk, on-site traffic accidents, has been tried to be visualized by the truck defined in the virtual environment, but this object was not sufficient in determining the risks (H3.7; $\rho = 0.250$) is intended to identify.

The hazard identification process of the trainees arises from mentally visualizing the construction task and evaluating the potential risks (Hadikusumo and Rowlinson, 2002). The indicators and other visual elements improve mental visualization process; thus, identifying risks become easier (Hadikusumo and Rowlinson, 2002). Quite consistently, Albert and colleagues (2014) developed a human-computer interactive augmented virtuality training platform was developed and cognitive mnemonics based on energy sources such as mechanical, electrical, etc. The observation of Albert and colleagues (2014) in developed augmented virtuality platform involve some similarities

with V-SAFE. Similar to V-SAFE, Albert and colleagues (2014) improved the trainees' mental visualization by defining triggers based on energy-sources. In this research, it is aimed to develop a similar but different approach. Rather than defining energy sources, it was aimed to integrate triggers by visualizing warnings and unsafe situations. As a result, the findings of the study reported by Albert and colleagues (2014) is quite similar to this research' results. The findings of the study showed that participants were able to identify a 46% of field risks before using an augmented virtuality-based safety training method, but they identified 77% of field risks after augmented virtuality use (Albert *et al.*, 2014). In other words, using the augmented virtuality-based safety training method improved the hazard identification level of trainees by 31%. The findings of Albert and colleagues (2014) are quite similar to this research. In the second group of the experiment, around 33% of the eight risk sources in total could be identified by conventional methods. Moreover, the risk identification level of trainees improved to 84.55%, when V-SAFE was used after conventional safety notes. This improvement corresponds to 51.55% improvement. The technology used by Albert and colleagues (2014) in occupational safety training methods showed some differences compared to this study. Albert and colleagues (2014) classified risk sources according to energy types. The data were transferred from building information modeling and construction scenarios. In short, it has transferred the real field data to the virtual environment using augmented virtuality. The occupational safety training method used in this study is similar because it is based on virtual environment. However, in Albert and colleagues' study (2014), risks are based on the principle of classifying over energy sources, rather than fatal accidents. Similarly, Sacks and colleagues (2013) tested the risk identification skills before, immediately after, and a month after the occupational safety training, thus assessing the effectiveness of occupational safety training in three different processes. The analysis results of Sacks and colleagues (2013) showed that virtual reality-based occupational safety training had a significant effect only in the immediately after process, and not any statistical impact was found in the analyzes, which took place one month later.In their research, Sacks and colleagues used a virtual reality-based occupational safety training program, and the experimental method used was similar to this study (2013). The main difference is that virtual reality uses a different representation method than the virtual environment. Similarly, a very similar

approach to experimental design was used by Dzeng and colleagues (2016). Dzeng and colleagues (2016) divided the participants into two groups as an experimental method and compared the virtual environment-based occupational safety training tool developed with conventional risk identification methods. The study results of Dzeng and colleagues (2016) showed that trainees' risk identification level was significantly more effective than traditional safety training methods under 99% level of significance. Similar to this study, Dzeng and colleagues (2016) associated risks to objects based on total 20 different risk sources. Unlike V-SAFE, the same threats were defined for more than one object in Dzeng and colleagues' study (2016). As an essential limitation of this research, hazardous situations are assigned to a single object in V-SAFE. To overcome these limitations, the same accident risks can be assigned to multiple objects in future versions of V-SAFE. In addition, the efficacy of V-SAFE in hazard identification process was measured immediately after the training, and its long-term effect was not examined. The long-term impact of V-SAFE can be measured in future studies.

To summarize, the effectiveness of the experimental method and virtual safety training method used in other studies (e.g., Albert *et al.*, 2014; Dzeng *et al.*, 2016; Sacks *et al.*, 2013) are similar to this study on risk identification. In addition, utilizing V-SAFE in improving the trainees' hazard identification level could be considered as statistically significant. Therefore, the following statement is accepted: Statement 3: *V-SAFE provides an effective hazard identification process for trainees.*

5.6. Hands-On Practice and Eye Tracking

Burke and colleagues (2006) highlight that an adequate safety training method should cover hands-on practice. The hands-on practice is vital in safety training process to reinforce safety knowledge (Burke *et al.*, 2006). To provide adequate hands-on practice, the previous studies (Hsiao *et al.*, 2014; Tichon and Diver, 2010) simulated some construction tasks through virtual environments. For instance, Tichon and Diver (2010) simulated plant operator activities by using an immersive virtual environment. The comparative analysis results show that the developed training method provides novice trainees an opportunity to gain adequate hands-on practice. Similarly, Hsiao and

colleagues (2014) executed an augmented virtuality-based training platform to improve the safety awareness of the trainees, during scaffolding activities. Hsiao and colleagues (2014) found that the developed model can provide adequate safety training during scaffolding tasks. Taken together, these studies highlight that virtual environments are able to provide hands-on practice.

To evaluate the effectiveness of hands-on practice, eye-tracking metrics could be considered as an appropriate approach. The hands-on training is defined as "In terms of a specific instructional strategy where trainees are actively engaged in manipulating materials' (Flick, 1993:9). As mentioned in previous sections, the hands-on practice should be executed under realistic working conditions and, the trainee should do the same practice, what they will be required to do every day (Sisson, 2001). As a result, the hands-on practice level could be evaluated through the safety behavior of trainees between the real and virtual construction site. Therefore, it could be concluded that if the safety behavior of trainees such as attention level and risk perception level are equally or more effective in V-SAFE compared to the real construction site, V-SAFE is able to provide adequate hands-on practice.

To determine whether V-SAFE provides an effective hands-on practice, a four-staged experiment was conducted with eleven construction workers. In each session, one worker entered the construction site and identified potential risk sources. Later on, the same worker entered the virtual environment of V-SAFE and asked to define potential risk sources, just like the first stage. During touring the real and virtual construction sites, primary eye-tracking metrics such as total fixation duration, first fixation duration, time to first fixation and fixation count were captured. In addition, each worker interacted with the instructor and identified the potential hazards by their verbal statements. Analysis results indicate that the majority of safety behavior such as attention, risk perception were preferable in V-SAFE environment compared to the real construction site.

As mentioned in section 3.1.4, the sensitivity to the potential risks by taking quick safety action is a crucial factor to prevent potential accidents (Wang *et al.*,

2016). In this sense, the eye-tracking behavior of trainees in terms of time to first fixation duration was analyzed. The analysis results show that the trainees' time to first fixation duration was significantly lower, compared to V-SAFE (Z = -4.18, $\rho <$ 0.01). On the construction sites, the trainees' lower time to the first fixation could be considered as an unfavorable situation, since, workers might not be able to take quick safety actions. In this sense, the visual search attitudes of workers on the real construction site were preferable, compared to V-SAFE. The major reason for higher time to first fixation duration in V-SAFE could be considered as the unsafe environment of real construction sites. For instance, a potential lack of attention on V-SAFE does not lead to an accident, due to risk-free nature of virtual environments. However, a potential error of a worker could lead to catastrophic accidents on real construction sites. Therefore, it was observed that the awareness and risk sensitivity of workers on the actual site was higher on the construction site, compared to V-SAFE. Therefore, following hypothesis is rejected: *Hypothesis 4.1: The time to first fixation durations of participants using V-SAFE are significantly lower than the time to first fixation durations of the same participants on the real construction site.*

The previous literature highlights that trainees' lack of attention, which leads to failure to define risky situations, is one of the main reasons of on-site accidents (Garrett and Teizer 2009; Rozenfeld *et al.* 2010). As a result of not identifying hazards accurately by trainees, they are not able to take correct safety action (Garrett and Teizer 2009; Rozenfeld *et al.* 2010). Therefore, a high level of attention is quite crucial in safety management context to identify hazards correctly. In order to compare, whether trainees' attention level is higher on V-SAFE training, compared to real construction site, their total fixation duration was analyzed. The analysis results show that the trainees' total fixation duration was significantly higher in V-SAFE environment, compared to the real construction site (Z = -3.99, $\rho < 0.01$). The analysis results provide a suitable outcome in terms of eye-tracking literature. The previous research (Garrett and Teizer 2009; Rozenfeld *et al.* 2010) states that the higher total fixation duration points out higher motivation, in other words, total fixation duration and attention level are directly proportional. Therefore, one can conclude that trainees' attention level was significantly higher in V-SAFE training and the following hypothesis is ac-

cepted: *Hypothesis 4.2: The total fixation durations of participants using V-SAFE are significantly higher than the total fixation durations of the same participants on the real construction site.*

Risk perception is another crucial parameter for hands-on practice effectiveness. According to Paek and Hove (2017:1), risk perception refers to "people's subjective judgments about the likelihood of negative occurrences such as injury, illness, disease, and death". In this sense, the risk perception is indispensable in the construction safety management context, since, it figures out which hazards individuals pay attention to and how they manage them (Paek and Hove, 2007). In their seminal study, Habibnezhad and colleagues (2016) analyzed the workers' risk perception impact to their visual search strategies. The analysis results show that the trainees with higher risk perception have higher first-fixation duration. In other words, the first fixation duration is an essential factor for the risk perception of trainees, and they are directly proportional. In this sense, higher first fixation duration is preferable in V-SAFE, since, higher first fixation duration proves the fact that V-SAFE improves the risk perception level of trainees. In order to check, whether trainees' risk perception level is higher on V-SAFE training, compared to real construction site, their first fixation duration was compared. The analysis results show that the trainees' first fixation duration was significantly higher in V-SAFE environment, compared to the real construction site ($Z = -3.99$, $\rho < 0.01$). Therefore, the following hypothesis is accepted: *Hypothesis 4.3: The first fixation durations of participants using V-SAFE are significantly higher than the first fixation durations of the same participants on the real construction site.*

As mentioned in the previous section, hazard identification is a crucial factor to provide on-site safety (Carter and Smith, 2006; Huang and Hinze, 2003; Patrucco *et al.*, 2010). Therefore, finding the parameters influence the efficiency of the hazard identification process is vital in construction industry. To fill this gap, Hasanzadeh and colleagues (2017) checked the impact of eye-tracking metrics of trainees to analyze the hazard identification ability based on the attentional level. The findings of Hasanzadeh and colleagues found that higher numbers of fixation count had the largest impact on the hazard identification performance of the trainees (2017). The analysis

results also indicate that fixation counts and the attention level of trainees are directly proportional (Hasanzadeh *et al.*, 2017). Therefore, higher fixation counts in V-SAFE provides evidence of higher attention level of trainees, compared to real construction sites. The analysis of this research shows that the number of trainees' fixation counts in V-SAFE are higher, compared to the number of fixations on the real construction site ($Z = -1.64$, $\rho < 0.05$). However, the analysis results also indicate that achieved power was 0.2903, which is quite lower than default power level. As a result of high probability of type II error, the results could be considered as trivial, and the following hypothesis is rejected: *Hypothesis 4.4: The fixation counts of participants using V-SAFE are significantly higher than the fixation count of the same participants on the real construction site.*

The analysis results of the research indicate that V-SAFE was able to improve the risk perception level of trainees. Moreover, the attention level of trainees was higher in V-SAFE, compared to the real construction site. Therefore, one could be concluded that major of the safety behavior parameters of trainees are more effective in V-SAFE compared to the actual construction site. Therefore, following statement is accepted: *Statement 4: Trainees' eye-tracking behavior in V-SAFE are more effective than the trainees on a real construction site.*

5.7. Gaming and Site Experience

As mentioned in the previous sections, construction companies have used information-based techniques such as lectures, videos, or presentations (Burke *et al.*, 2006). Consequently, construction workers fail to develop necessary implicit knowledge about the essentials of safety and gather adequate implicit knowledge through the tasks on the real construction site. While gathering required practical knowledge on the real construction site, the inexperienced worker could have potential severe accidents. Consistent with this statement, workers with less than 1-year of work experience are exposed to around 80% of all fatal accidents (Chi *et al.*, 2005).

To fill this gap, it is suggested that trainees without having any construction site experience should demonstrate a similar performance in V-SAFE training with the trainees having prior site experience. In other words, prior site experience should not be related to training performance in V-SAFE. Therefore, trainees without any construction site experience such as novice workers could be considered as the target population of virtual environment-based safety training. To assess the impact of the construction site on the effectiveness of V-SAFE training, two different study groups were compared. The first group involved students with a civil engineering background, while the subjects in the second group were from other disciplines such as computer engineering, mechanical engineering, etc. Briefly, subjects in the first group have site experience, since they completed site internship. In the second group of trainees do not have any construction site experience. Kruskal-Wallis H Test with Scheier-Ray-Hare Extension was applied to check, whether the performance of all trainees is similar to each other. The analysis results show that the training performance of the trainees with and without construction site experience is similar ($H = 0.045195701$, $\rho = 0.8316$). Therefore, trainees without site experience could benefit from the training provided by V-SAFE. This outcome has some significant practical implications. The previous analysis results of the repeatable virtual experiment (see Section 4.2) clearly indicate that V-SAFE is a suitable tool to provide necessary implicit knowledge through hands-on practice. In this analysis, it is also found that prior construction site experience does not impact the trainees' performance in V-SAFE. As stated, workers with less than 1-year work experience were exposed to around 80% of fatal accidents (Chi *et al.*, 2005). In this sense, the results offer a significant practical implication by a great opportunity for the novice trainees to provide sufficient safety training method. Instead of being exposed to on-site risks, novice workers could develop necessary implicit knowledge in a risk-free virtual environment. Therefore, following hypothesis is accepted: Hypothesis 5.1: Both trainees with and without site experience could use virtual construction safety training tool in a similar and effective manner.

Limited computer literacy of construction workers is another significant barrier for adequate safety through the virtual environment. The previous literature (e.g. De Kort and Ijsselsteijn, 2008; Nacke *et al.*, 2010) argue that the effectiveness of vir-

tual training methods is directly proportional to the computer gaming experience of trainees. In order to develop an effective safety training method, the user interface of the developed environment should be very simple; thus, trainees with limited computer literacy could use the environment effectively. To check, whether trainees with less gaming experience could also benefit from V-SAFE training, trainees grouped under four different gaming experience level. In addition, the trainees' safety training performance was compared through Kruskal-Wallis H Test with Scheier-Ray-Hare Extension. In other words, the results were evaluated, whether the performance of all trainees is similar to each other based on their gaming experience. The analysis results show that the training performance of all groups with varied gaming experience is similar to each other (H = 0.680702599, $\rho = 0.8777$). Therefore, the following hypothesis and statement are accepted:

- *Hypothesis 5.2: Both trainees with and without computer gaming experience could use virtual construction safety training tool in a similar and effective manner*
- *Statement 5: Trainees' construction site and gaming experience do not have any correlation with safety training performance.*

6. CONCLUSIONS

This study proposes a safety training measurement procedure for construction firms. In this respect, safety training effectiveness elements were adopted as the key parameters. The safety training effectiveness was evaluated through five major parameters as safety information transfer, instructional feedback mechanism, knowledge development, hazard identification, and hands-on practice. Data was collected through several methods as automated performance analysis procedure of developed virtual tool, the paper-based examinations results of trainees and the eye-tracking metrics gathered through Tobii eye-tracking devices. Besides, collected data was analyzed through varied data analysis methods based on experimental procedure and data types. Conducted data analysis methods involve Wilcoxon Signed Rank Test, One-Way ANOVA with Repeated Measures, McNemar Test, Mann-Whitney U Test, Kruskal Wallis-H Test with Scheirer-Ray-Hare extension to test the conducted hypotheses. Findings of the research reveal that V-SAFE is able to provide adequate safety training in terms of all evaluated parameters as safety information transfer, instructional feedback mechanism, knowledge development, hazard identification, and hands-on practice. Based on the research findings, the following recommendations are represented to improve safety training effectiveness. project management performance.

6.1. Key Findings of the Research

- Based on the investigation of trainees' answers in close-ended examination in information comprehensibility experiment, it was noticed that understandability of information is quite crucial. Construction companies opt for using conventional safety training methods such as lectures, videos, or pamphlets, which unable to represent the information clearly. During the investigation of answers, it was clearly seen that the majority of trainees were misunderstood the information provided by conventional methods. The similar misunderstanding of information could occur on real construction workers. In order to bridge this gap, it was found that virtual environment-based safety training methods significantly

reduce the misunderstandings. However, it was still not possible to consider virtual environment-based safety training methods as flawless materials about understandability. In other words, despite the virtual environment-based safety training methods are quite suitable to improve the understandability of information, it is not possible to state trainees' misunderstandings would be eliminated by using virtual environment-based safety training methods. Therefore, evaluating the trainees' understanding by other means such as paper-based, oral, or written examinations are also useful for construction safety management context.

- Based on the evaluation of interactions between the trainees in information comprehensibility experiment, it was found that trainees had problems in not being able to memorize information or inability to recall the information they had memorized. In terms of collaboration, the interaction between trainees might be considered as a desired situation; however, trainees sometimes gather wrong information from other trainees. Both lack of memorization or failing to recall information were arisen from lack of time to memorize information.In virtual environment-based safety training, it becomes possible to provide limitless time to check and memorize information via practice. However, the memorization level of trainees based on mnemonics should be evaluated by the safety management team.
- Knowledge development, and especially establishing a base for trainees to develop implicit knowledge, is a quite important process in safety management context. Conventional safety training methods fail to cover implicit knowledge due to no hands-on practice. As a result, trainees learn the essentials of safety during real on-site activities, and unfortunately, the novice trainees expose to serious accidents. In safety management context, it is quite crucial to provide hands-on practice for trainees to provide opportunities to develop implicit knowledge. In addition, evaluating their safety knowledge level through monitoring their safety attitudes and behavior on the real construction site is required to ensure their knowledge level. In order to check their knowledge level, as a common error, close-ended or written examinations should not be used, because these methods are merely based on safety information. Checking their attitudes and behavior by close supervision could be considered as the main element of understanding

their knowledge level. Therefore all construction safety management department should put attention for close supervision of workers' safety attitudes.

- Based on the evaluation of the results in the repeatable virtual experiment, it was found that trainees did not learn the basics of safety fully in the first attempt. To improve their learning, the risk-free and repetitive approach of virtual safety training can be seen as an indispensable element. In the first attempt, it was noticed that most trainees made some wrong decisions. Although trainees collect specific feedback based on their behavior, it will not be possible to measure how they react to this feedback unless a second attempt is made. In order to check whether the trainees have learned the information fully, it is necessary to evaluate their performance until the performance is perfect. In this sense, it is essential for the safety management departments to verify trainees' performance in the virtual environment before they enter the real construction site.
- Conventional safety training methods such as lectures, videos or safety notes fail to cover the dynamic construction tasks since they are quite limited to support trainees to visualize on-site activities mentally. Consequently, previous accident reports show that the majority of accidents are arisen from quite similar risk sources such as electrocution, working at height, structure collapses, etc. The majority of hazards could not be adequately identified by the workers. In order to bridge this gap, virtual environment-based safety training methods are quite suitable methods, because, they significantly enhance the mental visualization process of trainees by integrating the third dimension. However, visual elements used in the virtual environments are very substantial to determine the effectiveness of hazard identification process. In order to enhance the mental visualization of trainees, the quality of virtual elements has a high impact. In addition, virtual environments should also contain representation of dynamic tasks, hazardous materials, and risky situations to diversify the accident types. In this sense, construction companies should ensure that varied reasons for all accidents are visually supported by the virtual environment. In other words, they need to ensure that developed virtual environments should contain all precursors that lead to accidents. Consequently, the mental visualization process of workers could be improved and hazard identification process becomes efficient.

- Eye-tracking devices are quite suitable devices to measure the psychological elements of trainees. The majority of hazards arose from psychological factors such as attention level. In addition, crucial individual safety skills such as risk perception, sensitivity to potential risks, and quick safety action could be directly analyzed through several eye-tracking metrics. So, eye tracking technologies have high potential to evaluate safety training effectiveness, via integrated innovative algorithms to the safety behavior analysis of trainees. As mentioned in the previous sections that providing on-site safety is quite related to human-based factors such as, attention level, unsafe behavior, etc. In order to check human-based factors, it is quite essential for the construction safety management department to integrate eye-tracking devices into the safety training process.

6.2. Revisiting the Research Aim, Objectives and Major Contribution of Thesis

The main aim of the research was to quantitatively analyze the effectiveness of virtual construction safety training tool entitled Virtual Safety Analysis For Engineering applications (V-SAFE) in terms of all parameters of effective safety training. In order to achieve this aim, several objectives were developed as critical milestones of the thesis. The objectives of this research are as follows;

The first objective of the research was about measuring the effectiveness of virtual construction safety training method: V-SAFE. To achieve this objective, the previous literature was reviewed and the critical factors of effective safety training were determined. As a result, section 2.3 was developed, which constitutes the literature review upon this objective. After the completion of the review, the previous studies about virtual construction safety training were reviewed in section 2.4, in order to understand the different types of safety training effectiveness models used. After the completion of literature review, relevant hypotheses were developed, the experimental design of experiments were built and data analysis methods were assigned in chapter 3. Later on, collected datasets were analyzed and the analysis results were represented. As a result of the issues mentioned above, this objective is clearly accomplished.

The second objective of the research was about evaluating construction workers? behavior patterns through eye-tracking methods. To achieve this objective, the previous literature about eye-tracking studies used in construction safety training were reviewed in section 2.6. In addition, eye tracking metrics and their use areas in different studies related to psychology was investigated in 2.6.1. As a result, it becomes possible to bridge the eye-tracking metrics, human psychology and effective safety training parameters, which explained in detail in section 3.2.2. After the completion of the review, the experimental setting of eye movement behavior comparison experiment was set and the results of this analysis indicate that eye-tracking approach were quite successful to evaluate the trainees safety behavior in both virtual environment and on real construction site. As a result of the issues mentioned above, objective is clearly satisfied in the study.

The third objective of the research was about gathering relevant feedback about the use of V-SAFE from the workers; the main target group. In order to accomplish this objective, the eye-tracking experiments were conducted by using construction workers from two different construction companies, Bertuğ Bey Construction and Zetaş. In addition, demographic data of workers such as age, expertise, number of years of experience, etc. were collected through questionnaires. All collected data established a base for evaluating the effect of psychological factors such as behavior patterns, mental workload,etc. in the construction safety training process. As a result of the conducted analyses, it was objected to minimize the concerns about the use of students as subjects in construction health and safety training research. In brief, the objective is satisfied completely.

The fourth objective of the research was about developing a conceptual model that explains the effectiveness of virtual construction safety training methods. In order to accomplish this objective, the previous studies that focused on evaluating the effectiveness of the safety training methods were reviewed in section 2.3. In addition, the previous literature about the studies that focused on safety training and contained any features of mixed reality tools were reviewed in section 2.4. Based on the literature review, Table 2.4 and 2.5 were developed for systematically evaluating previous

studies and represent the major features of effective safety training. Later on, relevant hypotheses were developed, experimental design of experiments were built and data analysis methods were assigned in chapter 3. During the assignment of data analysis methods and experimental design, experience, implicit and implied knowledge was acquired. Through the experience and knowledge gained, it has been found out how the hidden effects of virtual security training methods can be evaluated in different parameters of effective security training. Based on acquired knowledge, a new conceptual model was created to evaluate all parameters of effective virtual safety training process (Fig 5.1). As a result of the issues mentioned above, the objective is clearly accomplished.

As the major contribution of the thesis, it is found that virtual tools are suitable to provide effective construction safety training. As could be seen from the key findings of the research and analysis results, all parameters of effective safety training could be provided through virtual construction safety training methods. As a major gap in the literature, the previous studies emphasized on only one or two of specific parameters of the efficient safety training. The findings of the previous studies offers some insights for the following studies to understand how virtual tools could be implemented to the safety training process. However, their methodology on evaluating safety training could be considered as quite limited, due to lack of focus on hands-on practice and behavioral modeling. So, it is also important to develop a conceptual model to integrate all parameters of virtual safety training and define key phases of evaluation of virtual safety training performance. In brief, this study offers new insights by integrating different knowledge areas as: virtual construction safety training, safety training effectiveness, eye tracking and psychology.

6.3. Limitations and Recommendations for Further Research

Although several contributions of the study explained in the previous section, there are also some limitations, which should be analyzed in the following research.

- Behavioral modeling, which is a quite important parameter for effective safety training, puts attention on the evaluation of trainees' safety behavior adjustment. So, the assessment of behavioral modeling could be seen as a vital part of effective safety training. In order to evaluate the impact of V-SAFE on behavioral modeling, it was required to assess the changes in trainees' behavior over a long period. Due to the time limitation of the research and impossibility to enter a real construction site daily, behavioral modeling could not be analyzed in the study. However, the previous literature (Burke *et al.*, 2006; Kirkpatrick, 1979) state that improving the knowledge level in training could lead to changes in the behavior of trainees. Therefore the adjustment of safety behavior was indirectly analyzed through knowledge development to minimize the limitation. Further research might explore the direct impact of virtual safety training methods to safety behavior adjustment process.
- Due to easy access to students, the conducted experiments were mainly used students as the subjects. For quite long time, the debate has still continued, whether students can be used as the surrogates of industry professionals in the engineering experiments. In other words, there are some disputes among the researchers about how well the results acquired from the students represent the results obtained by the industry professionals. To evaluate the validity of using students as surrogates, Runeson (2003) compared the performance of freshmen students, graduate students and industry professionals in the context of the personal software process. The analysis results indicate that quite similar improvements occurred between these three groups. Similarly, Salman and colleagues (2015) compared the results acquired from the students and industry professionals to comprehend, what level of students represent professionals as experimental subjects in software engineering. Salman and colleagues (2015) observed similar results between the students and industry professionals. In virtual safety training context, it was noticed that a considerable amount of previous studies (e.g., Bosché *et al.*, 2015; Chen *et al.*, 2016; Chi *et al.*, 2012; Ding *et al.*, 2013; Hadipriono and Barsoum, 2002; Juang *et al.*, 2013; Lu and Davis, 2016; Paes *et al.*, 2017; Pedro *et al.*, 2015; Perlman *et al.*, 2014; Teizer *et al.*, 2013; Le *et al.*, 2015) conducted experiments, while the students were the surrogates. Besides, a

study conducted by Perlman and colleagues (2014) aimed to compare the hazard identification level between the safety management team and students to evaluate, what extent students could be used as surrogates. Analysis results show that there are not any major differences between the performance of students and industry professionals. According to the high number of studies that used students as subjects in the virtual safety training context (e.g., Runeson, 2003; Salman *et al.*, 2015; Bosché *et al.*, 2015; Chen *et al.*, 2016; Chi *et al.*, 2012; Ding *et al.*, 2013; Hadipriono and Barsoum, 2002; Juang *et al.*, 2013; Lu and Davis, 2016; Paes *et al.*, 2017; Pedro *et al.*, 2015; Teizer *et al.*, 2013; Le *et al.*, 2015), and previously conducted analyses between students and industry professionals (e.g., Perlman *et al.*, 2014; Runeson, 2003; Salman *et al.*, 2015) one can conclude that students could be used during the experiments as surrogates. In this study, students were used as surrogates in four of the conducted experiments. However, the real construction workers were also used as subjects in eye-movement behavior experiment to overcome a potential limitation. Further studies might assess the outcome of virtual environment-based safety training methods using real construction workers as subjects.

- The mnemonics and memorization process were evaluated during the information comprehensibility experiment in this research. The results clearly indicated that V-SAFE was able to improve the memorization process of trainees by enhancing the mnemonics. However, due to the experimental procedure, the trainees' memory enhancement could be considered for long and short term impact. A further study could assess the long-term effects of developed virtual environments to the trainees' mnemonics
- To evaluate the improvement in knowledge development, it was necessary to analyze the changes in the behavior of trainees. However, due to the possibility of checking the safety behavior of trainees on a real construction site, the knowledge development was evaluated through monitoring the changes in the virtual environment. The analysis results significantly show that the knowledge level of trainees was improved. However, as a more eligible approach, further studies might determine the improvement in knowledge development through checking the changes the trainees on the real construction site.

6.4. ACKNOWLEDGMENT

This material is based in part upon work supported by Bogazici University Scientific Research Projects (BAP) under Grant No. 13202 and 7902. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Bogazici University Scientific Research Projects (BAP).

REFERENCES

Abdelhamid T.S. and J.G. Everett, "Identifying Root Causes of Construction Accidents", *Journal of Construction Engineering and Management*, Vol. 126, No. 1, pp. 52-60, 2000.

Albert A., M.R. Hallowell, B.C. Kleiner, A. Chen and M.J. Golparvar-Fard, "Enhancing Construction Hazard Recognition with High-Fidelity Augmented Virtuality", *Journal of Construction Engineering and Management*, Vol. 140, No. 7, pp. 1-11, 2014.

Albert A., M.R. Hallowell, M. Skaggs and B. Kleiner, "Empirical Measurement and Improvement of Hazard Recognition Skill", *Safety Science*, Vol. 93, No. 1, pp. 1-8, 2017.

Albright S.C. and W.L. Winston, *Business Analytics: Data Analysis and Decision Making*, 5th Edition, Nelson Education, 2014.

Al-Hakim L., *Information Quality Management: Theory and Applications*, 1st Edition, IGI Global, 2007.

Alliger G.M. and E.A. Janak "Kirkpatrick's Levels of Training Criteria: Thirty Years Later", *Personnel Psychology*, Vol. 42, No. 2, pp. 331-342, 1989.

Anders G., "Pilot's Attention Allocation During Approach and Landing: Eye and Head Tracking Research in Am A330 Full?ight Simulator", 11th *International Symposium on Avia-tion Psychology (ISAP)*, pp. 5-8, 2001.

Andersson B., J. Dahl, K. Holmqvist, J. Holsanova, V. Johansson, H. Karlsson and A. Wengelin, "Combining Keystroke Logging with Eye-Tracking", *Writing and Digital Media*, Vol. 17, No. 1, pp. 166-172, 2006.

Arthur Jr.W., Jr.W. Bennett, P.S. Edens and S.T. Bell, "Effectiveness of Training in Organizations: A Meta-Analysis of Design and Evaluation Features", *Journal of Applied Psychology*, Vol. 88, No. 2, pp. 234-245, 2003.

Azuma R., "A Survey of Augmented Reality", *Presence: Teleoperators and Virtual Environments*, Vol. 6, No. 4, pp. 355-385, 1997.

Bahn S. and L. Barratt-Pugh, "Safety Training Evaluation: The Case of Construction Induction Training and The Impact on Work-Related Injuries in The Western Australian Construction Sector", International Journal of Training Research, Vol. 12, No. 2, pp. 148-157, 2014.

Ballco P. and T. de-Magistris, "Attention and Choice: Use of Eye-Tracking in Choice Behavior Research for Nutritional Claims, An Empirical Evidence of Yogurts In Spain", *XV EAAE Congress "Towards Sustainable Agri-Food Systems: Balancing between Markets and Society"*, Parma, Italy. August 28 ? September, Vol. 1, 2017.

Banks J.S., J.S. Carson II, B.L. Nelson and D.M. *Nicol, Discrete-Event System Simulation*, 5th Edition, Prentice Hall, Upper Saddle River, Boston, USA, 2010.

Bartsch R.A. and K.M. Cobern, "Effectiveness of PowerPoint Presentations in Lectures", *Computers and Education*, Vol. 41, No. 1, pp. 77-86, 2003.

Behzadan A.H. and V.R. Kamat, "Enabling Discovery?Based Learning in Construction Using Telepresent Augmented Reality", *Automation in Construction*, Vol. 33, No. 1, pp. 3-10, 2013.

Benjaoran V. and S. Bhokha, "An Integrated Safety Management with Construction Management Using 4D CAD Model", *Safety Science*, Vol. 48, No. 3, pp. 395-403, 2010.

Bhoir S.A., S. Hasanzadeh, B. Esmaeili, M.D. Dodd and M.S. Fardhosseini, *Measuring*

Construction Workers" Attention Using Eye-Tracking Technology, 2015.

Bosché F., M. Abdel-Wahab and L. Carozza, "Towards A Mixed Reality System for Construction Trade Training", *Journal of Computing in Civil Engineering*, Vol. 30, No. 2, pp. 1-12, 2015.

Briesemeister J.K., *Managing Complex Construction Projects: A Systems Approach*, Auerbach Publications, Broken Sound Parkway NW, 2018.

Bureau of Labor Statistics (BLS), 2015 https://www.bls.gov/news.release/cfoi.t04.htm, accessed in April 2017.

Burke M.J., S.A. Sarpy, K. Smith-Crowe, S. Chan-Serafin, R.O. Salvador and G. Islam, "Relative Effectiveness of Worker Safety and Health Training Methods", *American Journal of Public Health*, Vol. 96, No. 2, pp. 315-324, 2006.

Büyüköztürk Ş., *Veri Analizi El Kitabı*, 24th Edition, Pegem A Yayıncılık, 2004.

Carrillo P., P. Chinowsky, "Exploiting Knowledge Management: The Engineering and Construction Perspective", *Journal of Management in Engineering*, Vol. 22, No. 1, pp. 2-10, 2006.

Carter G.S. and S.D. Smith, "Safety Hazard Identification on Construction Projects", *Journal of Construction Engineering: Management*, Vol. 132, No. 2, pp. 197-205, 2006.

Champion E., *Playing with the Past: Digital Games and the Simulation of History*, Springer, England, 2010.

ChanLin L.J., "Animation to Teach Students of Different Knowledge Levels", *Journal of Instructional Psychology*, Vol. 25, No. 3, pp. 166-175, 1998.

Chen A., M. Golparvar-Fard and B. Kleiner, "SAVES: A Safety Training Augmented Virtuality Environment for Construction Hazard Recognition and Severity Iden-

tification", *13th International Conference on Construction Applications of Virtual Reality*, London, UK, October pp. 30-31, 2013.

Chen Y.C., H.L. Chi, Kang, S. C. and Hsieh, S. H., "Attention-Based User Interface Design for A Tele-Operated Crane", *Journal of Computing in Civil Engineering*, Vol. 30, No. 3, pp. 1-12, 2016.

Cheng T. and J. Teizer, "Real-Time Resource Location Data Collection and Visualization Technology for Construction Safety and Activity Monitoring Applications", *Automation in Construction*, Vol. 34, No. 1, pp. 3-15, 2013.

Chi C.F., T.C. Chang and H.I. Ting, "Accident Patterns and Prevention Measures for Fatal Occupational Falls in the Construction Industry", *Applied Ergonomics*, Vol. 36 No. 4, pp. 391-400, 2005.

Chi H.L., Y.C. Chen, S.C. Kang and S.H. Hsieh, "Development of User Interface for Tele-Operated Cranes", *Advanced Engineering Informatics*, Vol. 26, No. 3, pp. 641-652, 2012.

Chi H.L., J. Wang, X. Wang, M. Truijens and P. Yung, "A Conceptual Framework of Quality-Assured Fabrication, Delivery and Installation Processes for Liquefied Natural Gas (LNG) Plant Construction", *Journal of Intelligent: Robotic Systems*, Vol. 79, No. 3, pp. 433-448, 2015.

Cohen A., M.J. Colligan, R. Sinclair, J. Newman and R.J. Schuler, *Assessing Occupational Safety and Health Training*, National Institute for Occupational Safety: Health, Cincinnati, Ohio, 1998.

Cohen J., *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition*, Routledge, New York, 2013.

Cohen J., *The Earth is Round (p<. 05)*, Routledge, NY, USA, 2016.

Cohen L.J., *The Implications of Induction, 4th Edition*, Routledge, 2019.

Cox D.R. and N.J.H. Small, "Testing Multivariate Normality", *Biometrika*, Vol. 65, No. 2, pp. 263-272, 1978.

Cuzick J., "A Wilcoxon?Type Test for Trend", *Statistics in Medicine*, Vol. 14, No. 4, pp. 445-446, 1995.

Coskun K. U. Ş. and I. Keskin, "Levene ve Bartlett Testleri Üzerine Bir İnceleme", *Selçuk Journal of Agriculture and Food Sciences*, Vol. 22, No. 44, pp. 78-83, 2008.

Daluba N.E., "Effect of Demonstration Method of Teaching on Students' Achievement in Agricultural Science", *World Journal of Education*, Vol. 3, No. 6, pp. 1-7, 2013.

De Kort Y.A.W. and W.A. Ijsselsteijn, "People, Places and Play: Player Experience in A Socio-Spatial Context", *Computers in Entertainment*, Vol. 6, No. 2, pp. 18, 2008.

DeLone W.H. and E.R. McLean, "Information Systems Success: The Quest for The Dependent Variable", *Information Systems Research*, Vol. 3, No. 1, pp. 60-95, 1992.

Demirkesen S. and D. Arditi, "Construction Safety Personnel's Perceptions of Safety Training Practices", *International Journal of Project Management*, Vol. 33, No. 5, pp. 1160-1169, 2015.

Derrick B., A. Ruck, D. Toher and P. White, "Tests for Equality of Variances Between Two Samples which Contain Both Paired Observations and Independent Observations", *Journal of Applied Quantitative Methods*, Vol. 13, No. 2, pp. 1-19, 2018.

Despotovic D., S. Cvetanovic, V. Nedic and M. Despotovic, "Economic, Social and Environmental Dimension of Sustainable Competitiveness of European Countries", *Journal of Environmental Planning: Management*, Vol. 59, No. 9, pp. 1656-1678,

2016.

Dester W.S. and D.I. Blockley, "Safety?Behaviour and Culture in Construction Engineering", *Construction and Architectural Management*, Vol. 2, No. 1, pp. 17-26, 1995.

Ding L.Y., C. Zhou, Q.X. Deng, H.B. Luo, X.W. Ye, Y.Q. Ni and P. Guo, "Real-Time Safety Early Warning System for Cross Passage Construction in Yangtze Riverbed Metro Tunnel Based on the Internet of Things", *Automation in Construction*, Vol. 36, No. 1, pp. 25-37, 2013.

Dolgunsöz E., *Identifying Eye Movement Behaviors and Discrepancies of Learners of English as A Foreign Language at Different Ielts General Reading Levels*, Ph.D. Thesis, Hacettepe Üniversitesi, 2015.

Duchowski A.T., "Eye Tracking Methodology", *Theory and Practice*, Vol. 328, No. 614, pp. 2-3, 2007.

Duchowski A.T., E. Medlin, A. Gramopadhye, B. Melloy and S. Nair, *Binocular Eye Tracking in VR for Visual Inspection Training*, Proceedings of the ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, November, pp. 8-10, 2001.

Dzeng R.J., C.T. Lin and Y.C. Fang, "Using Eye-Tracker to Compare Search Patterns Between Experienced and Novice Workers for Site Hazard Identification", *Safety Science*, Vol. 82, No. 1, pp. 56-67, 2016.

Egbu C.O. and H.S. Robinson, "Construction as A Knowledge-Based Industry", *Knowledge Management in Construction*, Vol. 4, pp. 31-49, 2005.

Ellis P.D., *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis and the Interpretation of Research Results*, Cambridge University Press, UK, 2010.

Ellis R., *Implicit and Explicit Learning, Knowledge and Instruction, Implicit, Explicit*

Knowledge in Second Language Learning, Testing and Teaching, USA, 2009.

Eraslan S., Y. Yeşilada and S. Harper, "Scanpath Trend Analysis on Web Pages: Clustering Eye Tracking Scanpaths", *Acm Transactions on the Web*, Vol. 10, No. 4, pp. 20-35, 2016.

European Commision, *The European Construction Sector: A Global Partner*, 2016, https://ec.europa.eu/growth/content/european-construction-sector-global-partner-0_en, accessed in 2019.

European Statistical Office (Eurostat), *Accidents at work statistics*, 2018b, https: ec.europa.eu/eurostat/statistics-explained/index.php?title=Accidents_at_work_statistics, accessed in 2019.

European Statistical Office (Eurostat), accessed in, 2018a. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hsw-n2-02andlang=en. accessed in 2019.

Fang Y. and Y.K. Cho, "Effectiveness Analysis from A Cognitive Perspective for A Real-Time Safety Assistance System for Mobile Crane Lifting Operations", *Journal of Construction Engineering and Management*, Vol. 143, No. 4, 2016.

Farnsworth B., *10 Most Used Eye Tracking Metrics and Terms*, 2018, https://imotions.com/blog/7-terms-metrics-eye-tracking/ accessed in 2019.

Field A., *Discovering Statistics Using IBM SPSS Statistics, 3rd Edition, Sage Publications*, Sussex, UK, 2009.

Fisher R.A., *Statistical Methods for Research Workers, Breakthroughs in Statistics*, Springer, NY, NY, USA, 1992.

Flick L.B., *The Meanings of Hands-On Science*, Routledge, USA, 1993.

Frese M. and D. Zapf, "Action as The Core of Work Psychology: A German Approach", *Handbook of Industrial and Organizational Psychology*, Vol. 4, No. 2, pp. 271-340,

1994.

Garrett J.W. and J. Teizer, "Human Factors Analysis Classification System Relating to Human Error Awareness Taxonomy in Construction Safety", *Journal of Construction Engineering: Management*, Vol. 135, No. 8, pp. 754-763, 2009.

Gerkes M., *Information Quality*', Paradox of the Web, 1997.

Ghasemi A. and S. Zahediasl, "Normality Tests for Statistical Analysis: A Guide for Non-Statisticians", *International Journal of Endocrinology and Metabolism*, Vol. 10, No. 2, pp. 486-489, 2012.

Girden E.R., *ANOVA: Repeated measures*, Sage, Newbury Park, CA, 1992.

Goh Y.M. and C.U. Ubeynarayana, "Construction Accident Narrative Classification: An Evaluation of Text Mining Techniques", *Accident Analysis and Prevention*, Vol. 108, No. 1, pp. 122-130, 2017.

Golovina O., J. Teizer and N. Pradhananga, "Heat Map Generation for Predictive Safety Planning: Preventing Struck-By and Near Miss Interactions Between Workers-On-Foot and Construction Equipment", *Automation in Construction*, Vol. 71, No. 1, pp. 99-115, 2016.

Goulding J., W. Nadim, P. Petridis and M. Alshawi "Construction Industry Offsite Production: A Virtual Reality Interactive Training Environment Prototype", *Advanced Engineering Informatics*, Vol. 26, No. 1, pp. 103-116, 2012.

Green S.B. and N.J. Salkind, *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*, Uppersaddle River: Prentice Hall, NJ, USA, 2008.

Greenhouse S.W. and S. Geisser, "On Methods in the Analysis of Profile Data", *Psychometrika*, Vol. 24, No. 2, pp. 95-112, 1959.

Gueorguieva R. and J.H. Krystal, "Move Over Anova: Progress in Analyzing Repeated-

Measures Data Andits Reflection in Papers Published in The Archives of General Psychiatry", *Archives of General Psychiatry*, Vol. 61, No. 3, pp. 310-317, 2004.

Guo H.L., H. Li and V. Li, "VP-Based Safety Management in Large-Scale Construction Projects: A Conceptual Framework", *Automation in Construction*, Vol. 34, No. 1, pp. 16-24, 2013.

Guo H., H. Li, G. Chan and M. Skitmore, "Using Game Technologies to Improve the Safety of Construction Plant Operations", *Accident Analysis and Prevention*, Vol. 48, No. 1, pp. 204-213, 2012.

Habibnezhad M., S. Fardhosseini, A.M. Vahed, B. Esmaeili and M.D. Dodd, *The Relationship Between Construction Workers' Risk Perception and Eye Movement in Hazard Identification*, Construction Research Congress 2016, San Juan, Puerto Rico, May 31 - June 2, 2016.

Hadikusumo B.H.W. and S. Rowlinson, "Capturing Safety Knowledge Using Design-For-Safety-Process Tool", *Journal of Construction Engineering and Management*, Vol. 130, No. 2, pp. 281-289, 2004.

Hadikusumo B.H.W. and S. Rowlinson, "Integration of Virtually Real Construction Model and Design-For-Safety-Process Database", *Automation in Construction*, Vol. 11, No. 5, pp. 501-509, 2002.

Hadipriono F.C. and A.S. Barsoum, "Modeling for Safety Against Falls from form Scaffolding in A Virtual Environment", *Civil Engineering Systems*, Vol. 19, No. 2, pp. 119-139, 2002.

Hale A., D. Walker, N. Walters and H. Bolt, "Developing The Understanding of Underlying Causes of Construction Fatal Accidents", *Safety Science*, Vol. 50, No. 10, pp. 2020-2027, 2012.

Han S., S. Lee and F. Peña-Mora, "Framework for A Resilience System in Safety

Management: A Simulation and Visualization Approach", *Computing in Civil and Building Engineering*, Proceedings of the International Conference, Anaheim, California, US, June, pp. 23-25, 2010.

Hasanzadeh S., B. Esmaeili and M.D. Dodd, "Measuring Construction Workers' Real-Time Situation Awareness Using Mobile Eye-Tracking", *In Construction Research Congress*, pp. 2894-2904, 2016.

Hasanzadeh S., B. Esmaeili and M.D. Dodd, "Impact of Construction Workers' Hazard Identification Skills on Their Visual Attention", *Journal of Construction Engineering and Management*, Vol. 143, No. 10, pp. 1-16, 2017.

Health and Safety Executive (HSE), *Construction Statistics in Great Britain, 2018*, 2018, http://www.hse.gov.uk/statistics/industry/construction.pdf. accessed in 2019.

Held E., C. Wolff, F. Gyntelberg and T. Agner, "Prevention of Work?Related Skin Problems in Student Auxiliary Nurses: An Intervention Study", *Contact Dermatitis*, Vol. 44, No. 5, pp. 297-303, 2001.

Henderson J.M., S.V. Shinkareva, J. Wang, S.G. Luke and J. Olejarczyk, "Predicting Cognitive State from Eye Movements", *Plos One*, Vol. 8, No. 5, pp. 1-6, 2013.

Heng L., D. Shuang, M. Skitmore, H. Qinghua and Y. Qin, "Intrusion Warning and Assessment Method for Site Safety Enhancement", *Safety Science*, Vol. 84, pp. 97-107, 2016.

Hesterberg T., D.S. Moore, S. Monaghan, A. Clipson and R. Epstein, "Bootstrap Methods and Permutation Tests", *Introduction to the Practice of Statistics*, Vol. 5, No. 1, pp. 1-70, 2005.

Hobday M., "Product Complexity, Innovation and Industrial Organisation", *Research Policy*, Vol. 26, No. 6, pp. 689-710, 1998.

Hoffman H., *Tactile Augmentation: Enhancing Presence in Virtual Reality with Tactile*

Feedback from Real Objects, Proceedings of Convention of the American Psychological Society, Toronto, Ontario, Canada, June 29-July 2, 1996.

Horng W.B., C.Y. Chen, Y. Chang and C.H. Fan, *Driver Fatigue Detection Based on Eye Tracking and Dynamic Template Matching*, IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan, March 21-23, 2004.

Hou L., H.L. Chi, W. Tarng, J. Chai, K. Panuwatwanich and X. Wang, "A Framework of Innovative Learning for Skill Development in Complex Operational Tasks", *Automation in Construction*, Vol. 83, pp. 29-40, 2017.

Howell D.C., *Statistical Methods for Psychology*, Cengage Learning, USA, 2009.

Hsiao H., P. Simeonov, B. Dotson, D. Ammons, T.Y. Kau and S. Chiou, "Human Responses to Augmented Virtual Scaffolding Models", *Ergonomics*, Vol. 48, No. 10, pp. 1223-1242, 2005.

Hua H., P. Krishnaswamy and J. Rolland, "Video-Based Eyetracking Methods and Algorithms in Head-Mounted Displays", *Optics Express*, Vol. 14, No. 10, pp. 4328-4350, 2006.

Huang K.T., Y.W. Lee and R.Y. Wang, *Quality Information and Knowledge*, Prentice Hall PTR, USA, 1998.

Huang X. and J. Hinze, "Analysis of Construction Worker Fall Accidents", *Journal of Construction Engineering and Management*, Vol. 129, No. 3, pp. 262-271, 2003.

Isaacowitz D.M., H.A. Wadlinger, D. Goren and H.R. Wilson, "Selective Preference in Visual Fixation Away from Negative Images in Old Age? An Eye-Tracking Study", *Psychology and Aging*, Vol. 21, No. 1, pp. 40-48, 2006.

Juang J.R., W.H. Hung and S.C. Kang, "Simcrane 3D+: A Crane Simulator with Kinesthetic and Stereoscopic Vision", *Advanced Engineering Informatics*, Vol. 27, No. 4, pp. 506-518, 2013.

Just M.A. and P.A. Carpenter, "Eye Fixations and Cognitive Processes", *Cognitive Psychology*, Vol. 8, No. 4, pp. 441-480, 1976.

Kayhani N., H. Taghaddos and M. Noghabaee, *Utilization of Virtual Reality Visualizations on Heavy Mobile Crane Planning for Modular Construction*, 2018 Proceedings of the 35th ISARC, Berlin, Germany, July 20-25, 2019.

Khan K., *Mnemonics and Study Tips for Medical Students*, CRC Press, USA, 2016.

Khoury H.M. and V.R. Kamat "High-Precision Identification of Contextual Information in Location-Aware Engineering Applications", *Advanced Engineering Informatics*, Vol. 23, No. 4, pp. 483-496, 2009.

Khushaba R.N., C. Wise, S. Kodagoda, J. Louviere, B.E. Kahn and C. Townsend, "Consumer Neuroscience: Assessing The Brain Response to Marketing Stimuli Using Electroencephalogram (EEG) And Eye Tracking", *Expert Systems with Applications*, Vol. 40, No. 9, pp. 3803-3812, 2013.

Kılıç S., "Etki Büyüklüğü", *Journal of Mood Disorders*, Vol. 4, No. 1, pp. 44-46, 2014.

Kıral I.A., S. Çomu and C. Kavaklıoğlu, *Enhancing the Construction Safety Training by Using Virtual Safety Training Tool: V-Safe*, Canadian Society for Civil Engineering International Construction Specialty Conference (ICSCCSCE), Vancouver, Canada, June 7-10, 2015.

Ki J.S., "Interactive Training Simulator for Aerial Working Platform in A Virtual Environment", *Computer Applications in Engineering Education*, Vol. 19, No. 4, pp. 733-738, 2011.

Kim H., C.R. Ahn and K. Yang "Identifying Safety Hazards Using Collective Bodily Responses of Workers", *Journal of Construction Engineering and Management*, Vol. 143, No. 2, 04016090, 2016.

Kirby A., V. Gebski and A.C. Keech, "Determining The Sample Size in A Clinical

Trial", *Medical Journal of Australia*, Vol. 177, No. 5, pp. 256-257, 2002.

Kirkpatrick D.L., "Techniques for Evaluating Training Programs", *Training: Development Journal*, Vol. 33, No. 6, pp. 78-92, 1979.

Klahr D., L.M. Triona and C. Williams, "Hands on What? The Relative Effectiveness of Physical Versus Virtual Materials in an Engineering Design Project by Middle School Children", *Journal of Research in Science Teaching*, Vol. 44, No. 1, pp. 183-203, 2007.

Krokos E., C. Plaisant and A. Varshney, *Spatial Mnemonics Using Virtual Reality*, in Proceedings of the 2018 10th International Conference on Computer and Automation Engineering pp. 27-30, ACM, 2018.

Lavado?Valenzuela R., J.M. Bravo, A.P. Junqueira?Kipnis, M. Ramos de Souza, C. Moreno, A. Alonso, T. Liberman?Kipnis, W.D. da Silva and A. Caballero, "Distribution of the HLA Class II Frequency Alleles in Patients with Leprosy from the Mid?West of Brazil", *International Journal of İmmunogenetics*, Vol. 38, No. 3, pp. 255-258, 2011.

Le Q.T., A. Pedro and C.S. Park, "A Social Virtual Reality Based Construction Safety Education System for Experiential Learning", *Journal of Intelligent and Robotic Systems*, Vol. 79, No. 3, pp. 487-506, 2015.

Lee K.P., H.S. Lee, M. Park, H. Kim and S. Han, "A Real-Time Location-Based Construction Labor Safety Management System‟", *Journal of Civil Engineering and Management*, Vol. 20, No. 5, pp. 724-736, 2014.

Lehman A., N. O'Rourke, L. Hatcher and E. Stepanski, *JMP for Basic Univariate and Multivariate Statistics: Methods for Researchers and Social Scientists*, 2nd Edition, Sas Institute, NC, USA, 2013.

Levene H., *Robust Tests for Equality of Variances*, Contributions to Probability and

Statistics, Essays in Honor of Harold Hotelling, pp. 279-292, 1961.

Li H., G. Chan and M. Skitmore, "Multiuser Virtual Safety Training System for Tower Crane Dismantlement", *Journal of Computing in Civil Engineering*, Vol. 26, No. 5, pp. 638-647, 2012.

Li H., G. Chan and M. Skitmore, "Visualizing Safety Assessment by Integrating the Use of Game Technology", *Automation in Construction*, Vol. 22, pp. 498-505, 2012.

Li H., M. Lu, G. Chan and M. Skitmore, "Proactive Training System for Safe and Efficient Precast Installation", *Automation in Construction*, Vol. 49, pp. 163-174, 2015.

Li H., G. Chan, T. Huang, M. Skitmore, T.Y. Tao, E. Luo and Y.F. Li "Chirp-Spread-Spectrum-Based Real Time Location System for Construction Safety Management: A Case Study", *Automation in Construction*, Vol. 55, pp. 58-65, 2015.

Lin K.Y., J.W. Son and E.M. Rojas, "A Pilot Study of A 3D Game Environment for Construction Safety Education", *Journal of Information Technology in Construction*, Vol. 16, No. 5, pp. 69-84, 2011.

Lingard H. and S. Rowlinson, *Occupational Health and Safety in Construction Project Management*, Routledge, USA, 2004.

Logitech, accessed in April , 2019a, https://www.logitechg.com/tr-tr/products/gamepads /extreme-3d-pro-joystick.html. accessed in August 2019.

Logitech, *Logitech F310*, 2019b, https://www.logitechg.com/tr-tr/products/gamepads /extreme-3d-pro-joystick.html. accessed in August 2019.

Lorge I., D. Fox, J. Davitz and M. Brenner, "A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance, 1920-1957", *Psychological Bulletin*, Vol. 55, No. 6, pp. 337-372, 1958.

Lu X. and S. Davis, "How Sounds Influence User Safety Decisions in A Virtual Construction Simulator", *Safety Science*, Vol. 86, pp. 184-194, 2016.

Magel R.C. and S.H. Wibowo, "Comparing The Powers of the Wald?Wolfowitz and Kolmogorov?Smirnov Tests", *Biometrical Journal*, Vol. 39, No. 6, pp. 665-675, 1997.

Maxwell S.E., H.D. Delaney and K. Kelley, *Designing Experiments and Analyzing Data: A Model Comparison Perspective*, Routledge, USA, 2017.

McCrum-Gardner E., "Which Is the Correct Statistical Test to Use?", *British Journal of Oral and Maxillofacial Surgery*, Vol. 46, No. 1, pp. 38-41, 2008.

McElduff F., M. Cortina-Borja, S.K. Chan and A. Wade, "When T-Tests or Wilcoxon-Mann-Whitney Tests Won't Do", *Advances in Physiology Education*, Vol. 34, No. 3, pp. 128-133, 2010.

Mello-Thoms C., C.F. Nodine and H.L. Kundel, *What Attracts the Eye to the Location of Missed and Reported Breast Cancers?*, Proceedings of The (2002) Symposium on Eye Tracking Research and Applications, New Orleans, Louisiana, March 28-30, 2002.

Milgram P. and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays", *Ieice Transactıons on Information and Systems*, Vol. 77, No. 12, pp. 1321-1329, 1994.

Mohd N.I., K.N. Ali, S. Bandi and F. Ismail, "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, 2019.

Muller K.E. and C.N. Barton, "Approximate Power for Repeated-Measures ANOVA Lacking Sphericity", *Journal of The American Statistical Association*, Vol. 84, No. 406, pp. 549-555, 1989.

Nachar N., "The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from The Same Distribution", *Tutorials in Quantitative Methods for Psychology*, Vol. 4, No. 1, pp. 13-20, 2008.

Nacke L.E., A. Drachen and S. Göbel, "Methods for Evaluating Gameplay Experience in A Serious Gaming Context", *International Journal of Computer Science in Sport*, Vol. 9, No. 2, pp. 1-12, 2010.

Nazim M. and B. Mukherjee, *Chapter 3 - Knowledge Management from a Library and Information Science Perspective*, Chandos Publishing, USA, 2016.

Nonaka I. and Nishiguchi T., *Knowledge Emergence: Social, Technical and Evolutionary Dimensions of Knowledge Creation*, Oxford University Press, USA, 2001.

Nonaka I., "A Dynamic Theory of Organizational Knowledge Creation", *Organization Science*, Vol. 5, No. 1, pp. 14-37, 1994.

Oxford Economics, *A Global Forecast for the Construction Industry to 2030*, 2015, https://www.pwc.fr/fr/assets/files/pdf/2016/01/pwc_global-construction-summit-2030-enr.pdf. accessed in August 2019.

Özdamar K., *Modern Bilimsel Araştırma Yöntemleri: Araştırma Planlama, Toplum ve Örnek Seçimi, Güç Analizi, Proje Hazırlama, Veri Toplama, Veri Analizi*, Bilimsel Rapor Yazımı, Kaan Kitabevi, İstanbul, Turkey, 2003.

Özorhon B., *Türkiye'de Inşaat Sektörü ve Dünyadaki Yeri*, İstanbul Ticaret Odası, İstanbul, Turkey, 2012.

Paek H.J. and T. Hove, *Risk Perceptions and Risk Characteristics*, Practical content, 2017.

Paes D., E. Arantes and J. Irizarry, "Immersive Environment for Improving the Understanding of Architectural 3D Models: Comparing User Spatial Perception Between Immersive and Traditional Virtual Reality Systems", *Automation in Construction*,

Vol. 84, pp. 292-303, 2017.

Park C.S. and H.J. Kim, "A Framework for Construction Safety Management and Visualization System", *Automation in Construction*, Vol. 33, pp. 95-103, 2013.

Patrucco M., D. Bersano, C. Cigna and F. Fissore, "Computer Image Generation for Job Simulation: An Effective Approach to Occupational Risk Analysis", *Safety Science*, Vol. 48, No. 4, pp. 508-516, 2010.

Patterson M.P., *Movement and Learning in Lecture Classroooms, PhD Dissertation*, University of Missouri, 2011.

Pedro A., Q.T. Le and C.S. Park, "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, pp. 1-10, 2015.

Perlman A., R. Sacks and R. Barak, "Hazard Recognition and Risk Perception in Construction", *Safety Science*, Vol. 64, pp. 22-31, 2014.

Pinheiro R.B.O., N. Pradhananga, R. Jianu and W. Orabi, *Eye-Tracking Technology for Construction Safety: A Feasibility Study*, ISARC, Proceedings of the International Symposium on Automation and Robotics in Construction, Auburn, AL, USA, July 18-21, 2016.

Pinto J.K. and P.W. Morris, *The Wiley Guide to Managing Projects*, John Wiley and Sons, USA, 2004.

Poole A. and L.J. Ball, *Eye Tracking in HCI and Usability Research*, IGI Global, 2016.

Razali N.M. and Y.B. Wah, "Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests", *Journal of Statistical Modeling and Analytics*, Vol. 2, No. 1, pp. 21-33, 2011.

Renz A., M. Solas, P.R. Almeida, M. Buhler, P. Gerbert, S. Castagnino and C. Roth-

baller, *Shaping the Future of Construction, A Breakthrough in Mindset and Technology*, World Economic Forum, USA, January, pp. 20-23, 2016.

Riffenburgh R.H., *Statistics in Medicine*, Academic Press, San Diego, CA, 1999.

Rozenfeld O., R. Sacks, Y. Rosenfeld and H. Baum, "Construction Job Safety Analysis", *Safety Science*, Vol. 48, No. 4, pp. 491-498, 2010.

Runeson P., *Using Students as Experiment Subjects-An Analysis on Graduate and Freshmen Student Data*, Proceedings of the 7th International Conference on Empirical Assessment in Software Engineering, NY, USA, September, pp. 20-21, 2003.

Rüppel U. and K. Schatz, "Designing A BIM-Based Serious Game for Fire Safety Evacuation Simulations", *Advanced engineering informatics*, Vol. 25, No. 4, pp. 600-611, 2011.

Sacks R., A. Perlman, and R. Barak, "Construction Safety Training Using Immersive Virtual Reality", *Construction Management and Economics*, Vol. 31, No. 9, pp. 1005-1017, 2013.

Salkind N.J., *Encyclopedia of Research Design*, Sage, USA, 2010.

Salman I., A.T. Mısırlı and N. Juristo, *Are Students Representatives of Professionals in Software Engineering Experiments?*, (2015) IEEE/ACM 37th IEEE International Conference on Software Engineering, Florence, Italy, May, pp. 16-24, 2015.

Scheirer C.J., W.S. Ray and N. Hare, "The Analysis of Ranked Data Derived from Completely Randomized Factorial Designs", *Biometrics*, Vol. 32, No. 2, pp. 429-434, 1976.

Schröeder R., "Defining Virtual Worlds and Virtual Environments", *Journal for Virtual Worlds Research*, Vol. 1, No. 1, 2008.

Schwab K., *The Global Competitiveness Report 2017-2018*, 2018, http://www3.weforum.org/docs/GCR2017-2018/05FullReport/TheGlobalCompetitivenessReport2017-%E2%80%932018.pdf, accessed in August 2019.

Scruggs T.E., M.A. Mastropieri, S.L. Berkeley and L. Marshak, "Mnemonic Strategies: Evidence-Based Practice and Practice-Based Evidence", *Intervention in School and Clinic*, Vol. 46, No. 2, pp. 79-86, 2010.

Sedlmeier P. and G. Gigerenzer, "Do Studies of Statistical Power Have an Effect on the Power of Studies?", *Psychological Bulletin*, Vol. 105, No. 2, pp. 309-317, 1989.

Shakhashiri B.Z., "Lecture Demonstrations", *Journal of Chemical Education*, Vol. 61, No. 11, pp. 10-10, 1984.

Shapira A. and A. Elbaz, "Tower Crane Cycle Times: Case Study of Remote-Control Versus Cab-Control Operation", *Journal of Construction Engineering and Management*, Vol. 140, No. 12, pp. 1-12, 2014.

Shapira A. and B. Lyachin, "Identification and Analysis of Factors Affecting Safety On Construction Sites with Tower Cranes", *Journal of Construction Engineering and Management*, Vol. 135, No. 1, pp. 24-33, 2009.

Shapiro S.S. and M.B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)", *Biometrika*, Vol. 52, No. 3, pp. 591-611, 1965.

Shen X. and E. Marks, "Near-Miss Information Visualization Tool in BIM for Construction Safety", *Journal of Construction Engineering and Management*, Vol. 142, No. 4, pp. 0401-5100, 2015.

Shepherd G.W., R.J. Kahler and J. Cross, "Crane Fatalities - A Taxonomic Analysis", *Safety Science*, Vol. 36, pp. 83-93, 2000.

Sherman W.R. and A.B. Craig, *Understanding Virtual Reality*, Morgan Kauffman, San Francisco, CA, 2003.

Sisson G.R., *Hands-On Training: A Simple and Effective Method for on the Job Training, Berrett-Koehler Publishers*, Littleton, Colorado, 2001.

Silva A.S. and D.M. Sutk, *Digital Cityscapes: Merging Digital and Urban Playspaces, 1st Edition, Peter Lang Inc.*, International Academic Publishers, USA, 2009.

Smith C.N., J.C. Frascino, R.O. Hopkins and L.R. Squire, "The Nature of Anterograde and Retrograde Memory Impairment After Damage to The Medial Temporal Lobe", *Neuropsychologia*, Vol. 51, No. 13, pp. 2709-2714, 2013.

Social Security Institution (SSI), *Causes of Fatal Occupational Accidents*, 2019b, http://www.guvenliinsaat.gov.tr/index.html, accessed in August 2019.

Social Security Institution (SSI), *Construction Sector Statistics 2017*, 2019a, http://www.guvenliinsaat.gov.tr/images/istatistik/%C4%B0statistik2017.pdf, accessed in August 2019.

Sokolowski J.A. and C.M. Banks, *Principles of Modeling and Simulation: A Multidisciplinary Approach*, John Wiley and Sons, NY, USA, 2011.

Squelch A.P., "Virtual Reality for Mine Safety Training in South Africa", *Journal of The Southern African Institute of Mining and Metallurgy*, Vol. 101, No. 4, pp. 209-216, 2001.

Laurent R. and P. Turk, "The Effects of Misconceptions on the Properties of Friedman's Test", *Communications in Statistics-Simulation and Computation*, Vol. 42, No. 7, pp. 1596-1615, 2013.

Sun Y., R. Fisher, F. Wang and H.M. Gomes, "A Computer Vision Model for Visual-Object-Based Attention and Eye Movements", *Computer Vision and Image Understanding*, Vol. 112, No. 2, pp. 126-142, 2008.

Szabo A. and N. Hastings, "Using IT in The Undergraduate Classroom: Should We Replace the Blackboard with PowerPoint?", *Computers and Education*, Vol. 35,

No. 3, pp. 175-187, 2000.

Taeger D. and S. Kuhnt, *Statistical Hypothesis Testing with SAS and R*, Chichester, UK, Wiley, pp. 219-236 2014.

Talmaki S. and V. Kamat, "Real-Time Hybrid Virtuality for Prevention of Excavation Related Utility Strikes", *Journal of Computing in Civil Engineering*, Vol. 28, No. 3, pp. 1-17, 2012.

Tam C.M., S.X. Zeng and Z.M. Deng, "Identifying Elements of Poor Construction Safety Management in China", *Safety Science*, Vol. 42, No. 7, pp. 569-586, 2004.

Tam V.W.Y. and I.W.H. Fung, "Tower Crane Safety in the Construction Industry: A Hong Kong Study", *Safety Science*, Vol. 49, No. 2, pp. 208-215, 2011.

Teizer J., T. Cheng and Y. Fang, "Location Tracking and Data Visualization Technology to Advance Construction Ironworkers' Education and Training in Safety and Productivity", *Automation in Construction*, Vol. 35, pp. 53-68, 2013.

Tichon J. and P. Diver, "Plant Operator Simulation: Benefits and Drawbacks for A Construction Training Organization", *Cognition, Technology and Work*, Vol. 12, No. 3, pp. 219-229, 2010.

Tobii Pro, *Product Description - Tobii X2-30 Eye Tracker, Tobii X2-60 Eye Tracker*, 2014, https://www.tobiipro.com/ siteassets/tobii-pro/product-descriptions/tobii-pro-x2-product-description.pdf, accessed in August 2019.

Tobii Pro, *Tobii Pro Glasses 2 - Product Description*, 2018, https://www.tobiipro.com/ siteassets/tobii-pro/product-descriptions/tobii-pro-glasses-2-product-description.pdf, accessed in August 2019.

Toker D., C. Conati and B.C.G. Steichen, *Individual User Characteristics and Information Visualization: Connecting The Dots Through Eye Tracking*, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atalanta,

USA, April 22-27, 2013.

Turkish Statistical Institution (TSI), *Branches of Economic Activity and Distribution of Employed Persons by Years, NACE Rev.2*, 2018b, http://tuik.gov.tr/PreIstatistik Meta.do?istab-id=2255,accessed in December 2018.

Turkish Statistical Institution (TSI), *Turnover by Economic Activities*, 2018a, http:// www.tuik. gov.tr/PreIstatistikTablo.do?istab-id=2777, accessed in December 2018.

U.S. *Bureau of Economic Analysis, Value Added by Industry*, 2018, https://apps.bea.gov/ iTable/iTable.cfm?ReqID=51andstep=1, accessed in December2018.

U.S. Census Bureau, *Number of Employees in The U.S. Construction Industry from 2000 to 2016 (in 1,000s)*, 2017, https://www.statista.com/statistics/187412/numberof-employees-in-us-construction/, accessed in August 2019

Unreal Engine Official Website, *Warner Bros. Interactive Entertainment Licenses Epic Games Unreal Engine 3*, 2011, https://www.unrealengine.com/en-US/blog/warner-bros-interactive-entertainment-licenses-epic-games-unreal-engine-3, accessed in Sep tember 2019.

Vacha-Haase T. and B. Thompson, "How to Estimate and Interpret Various Effect Sizes", *Journal of Counseling Psychology*, Vol. 51, No. 4, pp. 473-480, 2004.

Vasey M.W. and J.F. Thayer, "The Continuing Problem of False Positives in Repeated Measures ANOVA in Psychophysiology: A Multivariate Solution", *Psychophysiology*, Vol. 24, No. 4, pp. 479-486, 1987.

Vinodkumar M.N. and M. Bhasi, "Safety Management Practices and Safety Behaviour: Assessing The Mediating Role of Safety Knowledge and Motivation", *Accident Analysis and Prevention*, Vol. 42, No. 6, pp. 2082-2093, 2010.

Walpole R.E., H.R. Myers, L.S. Myers and K. Ye, *Probability and Statistics for Engineers and Scientists, Person Education, 9th Edition*, Pearson Education Interna-

tional, Upper Saddle River, NY, 2007.

Wang J., P.X.W. Zou and P.P. Li, "Critical Factors and Paths Influencing Construction Workers? Safety Risk Tolerances", *Accident Analysis: Prevention*, Vol. 93, pp. 267-279, 2016.

Wang R.Y., and Strong, D. M., "Beyond accuracy: What data quality means to data consumers", *Journal of Management Information Systems*, Vol. 12, No. 4, pp. 5-33, 1996.

Wang X. and P.S. Dunston, *System Evaluation of a Mixed Reality-Based Collaborative Prototype for Mechanical Design Review Collaboration*, International Conference on Computing in Civil Engineering, Cancun, Mexico, July 12-15, 2005.

Waxley K.N. and G.P. Latham, *Developing and Training Human Resources in Organization*, Prentice Hall, Upper Saddle River, NS, 2002.

Weaver K.F., V.C. Morales, S.L. Dunn, K. Godde and P.F. Weaver, *An Introduction to Statistical Analysis in Research: with Applications in The Biological and Life Sciences*, 1st Edition, John Wiley and Sons, USA, 2017.

Wilcoxon F., S.K. Katti and R.A. Wilcox, *Critical Values and Probability Levels for the Wilcoxon Rank Sum Test and the Wilcoxon Signed Rank Test*, Pearl River, New York, USA, 1970.

Wilkins J.R., "Construction Workers' Perceptions of Health and Safety Training Programmes", ", Vol. 29, No. 10, pp. 1017-1026, 2011.

Wilkinson L., "Statistical Methods in Psychology Journals: Guidelines and Explanations", *American Psychologist*, Vol. 54, No. 8, pp. 594-604, 1999.

Wirth O. and S.O. Sigurdsson, "When Workplace Safety Depends on Behavior Change: Topics for Behavioral Safety Research", *Journal of Safety Research*, Vol. 39, No. 6, pp. 589-598, 2008.

Wong K.W., P. Chan and K. Lo, *Factors Affecting the Safety Performance of Contractors and Construction Sites*, Proceedings of The Second International Conference of CIB Working Commission W99, Honolulu, Hawaii, March 24-27, 1999.

Yarbus A.L., *Eye Movements During Perception of Complex Objects*, Springer, Boston, MA, 1967.

Zhang S., F. Boukamp and J. Teizer, "Ontology-Based Semantic Modeling of Construction Safety Knowledge: Towards Automated Safety Planning for Job Hazard Analysis (JHA)", *Automation in Construction*, Vol. 52, pp. 29-41, 2015.

Zhao D. and J. Lucas, "Virtual Reality Simulation for Construction Safety Promotion", *International Journal of İnjury Control and Safety Promotion*, Vol. 22, No. 1, pp. 57-67, 2015.

Zhao D., A. McCoy, B. Kleiner and Y. Feng, "Integrating Safety Culture into OSH Risk Mitigation: A Pilot Study on the Electrical Safety", *Journal of Civil Engineering and Management*, Vol. 22, No. 6, pp. 800-807, 2016.

Zhou Y., L.Y. Ding and L.J. Chen, "Application of 4D Visualization Technology for Safety Management in Metro Construction", *Automation in Construction*, Vol. 34, pp. 25-36, 2013.

APPENDIX A: THE EYE MOVEMENT BEHAVIOR EXPERIMENT RESULTS

Area of Interest	Environment	Time To First Fixation	Total Fixation Duration	Fixation Count	First Fixation Duration
1	Real Construction Site	1.4090	1.4417	7	0.2299
2	Real Construction Site	0.0300	0.5704	5	0.0208
3	Real Construction Site	60.2660	0.3207	4	0.0908
4	Real Construction Site	0.9790	0.4195	6	0.0431
5	Real Construction Site	0.3700	0.2437	2	0.1799
6	Real Construction Site	0.0000	0.2817	2	0.1818
7	Real Construction Site	9.6650	0.0546	1	0.0546
8	Real Construction Site	0.6790	2.3580	13	0.2081
9	Real Construction Site	0.0200	6.9115	35	0.1200
10	Real Construction Site	0.2000	0.5450	3	0.0105
11	Real Construction Site	0.5890	2.8713	15	0.1799
12	Real Construction Site	3.9470	0.2745	1	0.2745
13	Real Construction Site	3.9880	0.3696	2	0.3398
14	Real Construction Site	0.0100	2.3664	11	0.1099
15	Real Construction Site	4.7470	2.5536	16	0.3098
16	Real Construction Site	1.9090	0.0405	1	0.0405
17	Real Construction Site	0.0100	5.3079	21	0.0528
18	Real Construction Site	6.4170	0.1299	1	0.1299
19	Real Construction Site	1.5290	2.6094	11	0.0700
20	Real Construction Site	0.5200	3.1807	11	0.2225
21	Real Construction Site	0.0800	1.6702	12	0.0799
22	Real Construction Site	7.7960	1.2106	7	0.1590
23	Real Construction Site	0.0100	1.3653	11	0.1099
24	Real Construction Site	1.1600	0.1499	1	0.1499
25	Real Construction Site	1.3790	1.5405	11	0.0356
26	Real Construction Site	1.4790	1.5324	10	0.0295
27	Real Construction Site	0.1800	0.2099	2	0.1000
28	Real Construction Site	4.1980	0.1191	2	0.0454
29	Real Construction Site	0.2200	2.6497	18	0.6503
30	Real Construction Site	0.0000	0.6196	3	0.1099
31	Real Construction Site	0.1200	2.0626	9	0.1520
32	Real Construction Site	0.0000	7.1111	40	0.1653
33	Real Construction Site	0.0000	1.9814	8	0.0925
34	Real Construction Site	0.0000	0.0449	1	0.0449
35	Real Construction Site	1.6190	0.4897	1	0.4897
36	Real Construction Site	0.0290	4.6311	28	0.0049
37	Real Construction Site	0.1700	5.6059	20	0.1499
38	Real Construction Site	2.3090	2.2085	13	0.0567
39	Real Construction Site	1.2390	0.4297	2	0.3598
40	Real Construction Site	0.3900	0.0508	1	0.0508
41	Real Construction Site	0.0600	6.5598	35	0.2764
42	Real Construction Site	0.0000	0.3244	3	0.1545
43	Real Construction Site	0.8590	0.3298	1	0.3298
44	Real Construction Site	1.7590	0.2968	3	0.1599
45	Real Construction Site	0.2700	0.2694	4	0.0724

Figure A.1. The Eye Movement Behavior Experiment Results 1.

Area of Interest	Environment	Time To First Fixation	Total Fixation Duration	Fixation Count	First Fixation Duration
46	Real Construction Site	5.5270	0.2599	2	0.1799
47	Real Construction Site	7.0960	0.5697	2	0.1299
48	Real Construction Site	0.2200	1.9597	14	0.3098
49	Real Construction Site	0.9600	0.7709	8	0.0296
50	Real Construction Site	0.9100	0.1005	1	0.1005
51	Real Construction Site	1.6790	0.1013	1	0.1013
52	Real Construction Site	6.0160	0.1299	1	0.1299
53	Real Construction Site	9.3440	0.5361	3	0.0664
54	Real Construction Site	5.5960	0.0163	1	0.0163
55	Real Construction Site	5.7960	0.0165	2	0.0103
56	Real Construction Site	5.1870	0.1417	2	0.0317
57	Real Construction Site	13.8110	9.7754	41	0.4497
58	Real Construction Site	12.0430	0.4832	3	0.1734
59	Real Construction Site	2.3080	1.3957	6	0.2246
60	Real Construction Site	5.1170	1.8345	5	0.0387
61	Real Construction Site	22.0370	5.7911	16	0.2495
62	Real Construction Site	15.4410	6.0919	19	0.0030
63	Real Construction Site	6.4270	0.0198	1	0.0198
64	Real Construction Site	12.3430	1.1235	6	0.1699
65	Real Construction Site	0.0000	0.0483	1	0.0483
66	Real Construction Site	11.7040	1.8067	5	0.4697
67	Real Construction Site	0.9790	0.2540	2	0.1699
68	Real Construction Site	60.6250	1.9893	9	0.2499
69	Real Construction Site	4.1080	0.1899	1	0.1899
70	Real Construction Site	11.1940	9.8796	28	0.1922
71	Real Construction Site	18.7600	0.4087	2	0.0689
72	Real Construction Site	0.0500	6.1961	23	0.0333
73	Real Construction Site	1.4000	0.3898	1	0.3898
74	Real Construction Site	23.3370	4.2224	7	0.2033
75	Real Construction Site	2.6580	0.5197	4	0.2299
76	Real Construction Site	27.1850	0.5497	1	0.5497
77	Real Construction Site	0.0000	0.4529	2	0.1531
78	Real Construction Site	0.9790	9.4576	37	0.2386
79	Real Construction Site	6.7360	1.6253	11	0.2099
80	Real Construction Site	24.1860	18.9009	63	0.2199
81	Real Construction Site	29.3530	0.8471	10	0.0180
82	Real Construction Site	0.0000	0.1671	3	0.0878
83	Real Construction Site	0.0000	1.7916	7	0.1226
84	Real Construction Site	0.1590	0.3988	5	0.0800
85	Real Construction Site	32.6210	0.2167	2	0.0687
86	Real Construction Site	0.0000	0.4738	3	0.1140
87	Real Construction Site	28.3540	0.0668	1	0.0668
88	Real Construction Site	11.6830	8.0253	21	0.0999
89	Real Construction Site	0.0000	0.5242	6	0.0366
90	Real Construction Site	8.1050	0.1755	1	0.1755
91	Real Construction Site	6.5160	0.2806	3	0.1199

Figure A.2. The Eye Movement Behavior Experiment Results 2.

Area of Interest	Environment	Time To First Fixation	Total Fixation Duration	Fixation Count	First Fixation Duration
92	V-SAFE	14.3130	0.2400	1	0.2400
93	V-SAFE	1.5810	16.7900	63	0.1800
94	V-SAFE	3.3440	0.8500	4	0.5000
95	V-SAFE	57.9240	0.9600	4	0.0700
96	V-SAFE	29.0820	0.4300	1	0.4300
97	V-SAFE	25.8910	0.6000	2	0.2500
98	V-SAFE	7.2200	2.2800	8	0.2300
99	V-SAFE	22.5400	2.4000	9	0.0300
100	V-SAFE	22.5400	2.3100	7	0.0300
101	V-SAFE	0.0000	5.5800	19	0.1100
102	V-SAFE	0.0000	14.0200	41	0.4000
103	V-SAFE	0.0000	10.6000	29	0.3100
104	V-SAFE	0.0000	3.1400	11	1.0500
105	V-SAFE	0.0920	0.8100	2	0.7400
106	V-SAFE	6.4250	7.7400	21	2.3000
107	V-SAFE	55.9740	0.5800	3	0.2400
108	V-SAFE	84.8380	1.0000	4	0.2300
109	V-SAFE	0.1570	1.1000	3	0.3000
110	V-SAFE	0.5800	5.1800	11	0.1900
111	V-SAFE	0.7950	5.0100	9	2.3000
112	V-SAFE	97.4670	4.0900	15	0.0800
113	V-SAFE	127.6620	6.2400	20	0.3300
114	V-SAFE	102.4820	5.1600	22	0.1400
115	V-SAFE	1.0890	0.2200	1	0.2200
116	V-SAFE	0.2820	13.4400	48	0.3000
117	V-SAFE	7.8740	0.5400	1	0.5400
118	V-SAFE	0.8090	0.4900	2	0.3400
119	V-SAFE	13.7530	0.0800	1	0.0800
120	V-SAFE	1.7940	0.5300	4	0.1000
121	V-SAFE	10.0300	0.3700	3	0.2900
122	V-SAFE	10.0300	0.7500	5	0.2900
123	V-SAFE	0.2620	3.6400	18	0.1800
124	V-SAFE	1.2560	9.3300	35	0.5500
125	V-SAFE	0.4420	0.0800	1	0.0800
126	V-SAFE	28.5550	2.1100	9	0.3200
127	V-SAFE	0.5310	2.4900	9	0.1600

Figure A.3. The Eye Movement Behavior Experiment Results 3.

APPENDIX B: THE QUESTION SHEET OF CRANE OPERATOR ROLE

1) Which of the following personal protective equipment should be taken by the crane operator? Mark the equipment(s) that must be taken (☑).

☐ Steel Toe Boots ☐ Hardhat ☐ Reflective Vest ☐ Protective Gloves ☐ Safety Belt
☐ Goggles ☐ Walky-Talky

2) Which of the following is <u>not</u> one of the primary responsibilities of the crane operator?

a) To activate the warning system, before the load starts to move and during the duration of the movement.
b) To check the movements of the tower crane without the load, before proceeding with lifting.
c) To receive guidance from the pointer regarding the dead spots in the field of view.
d) To choose the rope according to the dimensions of the load to be lifted (width, length, height).

3) Which rope should be preferred for a load given the following size, weight, and wind strength information?

Wind Intensity: 23 km/h
Load Weight: 4.8 ton
Width: 2 meters
Length: 3.6 meters
Height: 2.2 meters

a) Type 1
b) Type 2
c) Type 3

4) Which of the load and environment information is <u>not</u> available on the crane main screen?

a) Wind Intensity
b) Temperature
c) Load Dimensions and Load Weights

Figure B.1. The Question Sheet of Crane Operator Role 1.

a) Forward
b) Down
c) Right
d) Side

6) During the loading of the material, which of the following operation(s) is/are the responsibility of the crane operator?

I. Lifting the load for measuring the equilibrium point of the material.
II. Ensuring that the hook on the tower crane and the connection equipment at the end touch the load.
III. Warning of the pointer and bricklayer on the site not to pass under the load.

a) Only I **b)** I and II **c)** II and III **d)** Only III **e)** I, II and III

7) Which of the following options or crane operator should pay attention to the transport of the load is given together?

I. Lifting the load for measuring the equilibrium point of the material.
II. Ensuring that the hook on the tower crane and the connection equipment at the end touch the load.
III. Warning of the pointer and brick masonry worker on the site not to pass under the load.

a) Only I **b)** Only II **c)** Only III **d)** I and II **e)** I, II and III

8) How should the crane operator provide the weight information of the material to be loaded?

a) Load information can be displayed on the tower crane main screen.
b) Display load information by capturing the correct angle on the field.
c) The load information is learned by the bricklayer.
d) Load information before the tower crane, by going to the load can learn.

Figure B.2. The Question Sheet of Crane Operator Role 2.

APPENDIX C: THE QUESTION SHEET OF POINTER ROLE

1) Which of the following personal protective equipment should be taken by the pointer? Mark the equipment(s) that must be taken (☑).

☐ Steel Toe Boots ☐ Hardhat ☐ Reflective Vest ☐ Protective Gloves ☐ Safety Belt
☐ Goggles ☐ Walky-Talky

2) Which of the following is **not** one of the primary responsibilities of the pointer?

a) To confirm that the load is ready for lifting.
b) To inform the tower crane operator about the lifting points of the load.
c) To guide the tower crane operator on the blind spots, not in view.
d) To attach himself to the column with the safety belt during unloading process.

3) Which of the following commands should be given for the movement of the trolley?

a) Forward - Backward
b) Up – Down
c) Right - Left

4) Which of the following statement(s) is **not** one of the tasks of the pointer?

I. To obtain wind intensity information from a bricklayer.
II. To check whether the hook and the connection point on end touch the load.
III. To determine the installation points.

a) All of these are the duties of the pointer
b) Only I
c) Only II
d) Only III
e) I and III

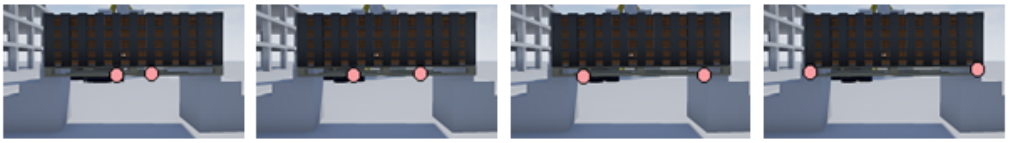
5) What is the principle of loading the load in V-SAFE?

a) Six-point of support principle
b) Six-point of hinge principle
c) Four-point of support principle
d) Four-point of hinge principle

Figure C.1. The Question Sheet of Pointer Role 1.

6) Which of the following loads can be considered more accurate as material loading?

a) **b)** **c)** **d)**



7) What issue(s) should the pointer pay attention to when loading the material?

I. Wind intensity
II. Weight of load
III. Crane operator commands
IV. The load center of gravity and loading points

a) I and III **b)** I and IV **c)** II and III **d)** Only III **e)** Only IV

8) During the lowering of the load, which or which of the following are the duties of the pointer?

I. To attach a safety belt
II. To guide the crane operator
III. To support the bricklayer during unloading of the load.
IV. To determine the safe point where the load is to be downloaded and notify the bricklayer.

a) I, III and IV
b) I and IV
c) II and III
d) II and IV
e) All of these are the duties of the pointer

Figure C.2. The Question Sheet of Pointer Role 2.

APPENDIX D: THE QUESTION SHEET OF BRICKLAYER ROLE

1) Which of the following personal protective equipment should be taken by the pointer? Mark the equipment(s) that must be taken (☑).

☐ Steel Toe Boots ☐ Hardhat ☐ Reflective Vest ☐ Protective Gloves ☐ Safety Belt
☐ Goggles ☐ Walky-Talky

2) Which of the following is not one of the primary responsibilities of the bricklayer?

a) To confirm that the load is ready for lifting.
b) To determine the lifting points of the load.
c) To inform the tower crane operator about the size and weight of the load.
d) To attach itself to the column with the belt apparatus during the lowering of the load.

3) Which of the following information should be notified to the tower crane operator by the bricklayer?

I. Wind Intensity
II. Temperature
III. Load Size
IV. Load Weight

a) I, III and IV **b)** Only IV **c)** III and IV **d)** II and III **e)** II, III and IV

4) Which of the following(s) is not one of the tasks of the bricklayer?

I. To inform the tower crane operator of load information.
II. To check whether the hook and the connection point on end touch the load.
III. To determine the loading points.

a) All of these are the tasks of the bricklayer.
b) Only I
c) Only II
d) Only III
e) II and III

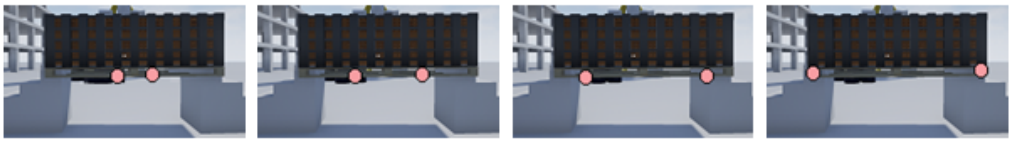
Figure D.1. The Question Sheet of Bricklayer Role 1.

5) What is the principle of loading the load in V-SAFE?

a) Six-point of support principle
b) Six-point of hinge principle
c) Four-point of support principle
d) Four-point of hinge principle

6) Which of the following loads can be considered more accurate as material loading?

a) **b)** **c)** **d)**



7) What aspect(s) should the bricklayer pay attention during the loading of the material?

I. Wind intensity
II. Whether the hook touches the load
III. Tower Crane Operator commands
IV. The load center of gravity and loading points

a) I and III **b)** I and IV **c)** II and III **d)** Only II **e)** Only IV

8) Which of the following responsibilities belong to the bricklayer during the unloading of the load?

I. Fixing the safety belt
II. Directing the Crane Operator
III. To determine the safe lowering point during the unloading of the load.

a) Only I **b)** I and II **c)** Only II **d)** II and III **e)** I and III

Figure D.2. The Question Sheet of Bricklayer Role 2.