Towards a Customizable Immersive Virtual Reality Serious Game for Indoor Earthquake Emergency Training

Zhenan Feng

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Civil Engineering, 2020.

Main Supervisor: Dr. Vicente A. González

Co-supervisors: Prof. Robert Amor

Prof. Carol Mutch

Department of Civil and Environmental Engineering

Faculty of Engineering

The University of Auckland

Auckland, New Zealand

Abstract

Enhancing the earthquake behavioral response and post-earthquake evacuation preparedness of building occupants is beneficial to increasing their chances of survival and reducing injured people and casualties after the mainshock of an earthquake. Traditionally, training approaches such as seminars, posters, videos or drills are applied to enhance preparedness. However, they are not highly engaging and have limited sensory capabilities to mimic life-threatening scenarios for the purpose of training. Immersive Virtual Reality (IVR) and Serious Games (SGs) as innovative digital technologies can be used to create training tools to overcome these limitations. SGs are an innovative approach devoted to training and educating people in a gaming environment. IVR is a technology to fully immerse users in a computer-generated virtual environment. Recently, increasing attention has been paid to IVR-based SGs for evacuation training to promote cognitive learning with highly credible experience. IVR SGs have been introduced to train individuals in specific building layouts or settings with fixed training objectives. However, the lack of flexibility in existing IVR SG's frameworks makes it challenging to have widespread uptake as trainees require different training objectives, pedagogical strategies, context, and content. As a result, the effectiveness of IVR SG-based training is jeopardized if the customization ability is limited.

In order to address the limitations mentioned above, this research proposes a customization framework for IVR SGs suited to earthquakes and post-earthquake evacuation training. A general framework for the effective development and implementation of IVR SGs suited to evacuation training was constructed based on a systematic literature review. Following that, an experiment taken at Auckland City Hospital was conducted to validate the general framework. Based on the general framework, we establish a customization framework incorporating the concept of adaptive game-based learning. According to this customization framework, trainees can receive earthquake emergency training in context by customizing virtual environments, storylines, and teaching methods.

Two experiments were carried out to provide empirical evidence to validate the customization framework, with one targeting children in a junior secondary school setting and the other targeting adults in an office setting. The usability results from both experiments showed that the customization process was easy to carry out, and the training experience with the customizable IVR SGs was adaptive and easy to follow for optimum learning. The school experiment also investigated the effectiveness of the teaching methods provided in the customization framework. Results suggested that post-game assessment and prior instruction were both effective for children to enhance knowledge and self-efficacy, with the integration of IVR SGs. The office experiment also investigated the effectiveness of the storytelling methods provided in the customization framework. Results revealed that both spiral narratives and linear narratives were effective to deliver the knowledge about behavioral responses in earthquakes and post-earthquake evacuation, with the integration of immediate feedback in IVR SGs.

In conclusion, a customization framework for IVR SGs suited to earthquake emergency training has been proposed and validated by this research, contributing to the knowledge of IVR SGs-based emergency training. Future research is suggested to extend the customization framework to including other types of emergency situations, and system-controlled adaptation (i.e., dynamic adaptation throughout training with the use of artificial intelligence).

Dedication

To Yi, my beloved wife

To Junyi, my beloved son

To Jianwei and Yaqing, my beloved parents

Acknowledgement

I have been lucky to be accompanied by a group of wonderful people along my Ph.D. journey. I am extremely grateful to my main supervisor, Dr. Vicente A. González, for his farseeing vision, insightful advice, and endless support like an old friend. I would also like to express my greatest gratitude to my co-supervisors Prof. Robert Amor and Prof. Carol Mutch, for their inspiring guidance and creative ideas. It would be a fairy tale for me to finish my Ph.D. program without the patience and support from my supervisors.

I wish to extend my appreciation to Dr. Guillermo Cabrera-Guerrero (Pontificia Universidad Católica de Valparaíso), Prof. Nan Li (Tsinghua University), Anouar Baghouz (University of Mons), Anass Rahouti (University of Mons), Bruce Rooke (Holonize), Yifan Gao (University of Auckland), and Mohammed Adel Abdelmegid (University of Auckland), Dr. Michael Spearpoint (OFR consultants), Prof. Rafael Sacks (Technion – Israel Institute of Technology), Dr. Jared Thomas (WSP Opus) for collaboration, feedback, support, and encouragement during my Ph.D. journey.

My heartfelt gratitude also goes to my beloved family. Their understanding, support, encouragement, and endless love are like a lighthouse to a sailor in a stormy sea.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CAVE	Cave Automatic Virtual Environment	IVR	Immersive Virtual Reality
FPS	Frames Per Second	NPC	Non-player Character
GBL	Game-based Learning	PBG	Problem-based Gaming
HMD	Head-mounted Display	SGs	Serious Games

List of Publications and Co-Authorship Forms

The contents of six chapters of this thesis are based on five research papers, published or submitted for publication on internationally recognized peer-reviewed journals. Apart from the changes to the formatting and minor changes in the text, the manuscripts appear as they have been published or submitted for publication.

Co-Authorship Form A

Section 3.2 A general framework for IVR SG-based emergency studies

The content of this section is extracted from:

Feng, Z., González, V.A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive Virtual Reality Serious Games for Evacuation Training and Research: A Systematic Literature Review. *Computers & Education, 127*, 252-266.

Co-Authorship Form B

Chapter 4: An IVR SG for Earthquake Emergency Training

The content of this chapter is extracted from:

Feng, Z., González, V.A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., Lovreglio, R., & Cabrera-Guerrero, G. (2020). An Immersive Virtual Reality Serious Game to Enhance Earthquake Behavioral Responses and Post-earthquake Evacuation Preparedness in Buildings. *Advanced Engineering Informatics, 45*, 101118.

Co-Authorship Form C

Chapter 5: A Customization Framework for IVR SGs Suited to Earthquake Emergency Training

The content of this chapter is extracted from:

Feng, Z., González, V.A., Mutch, C., Amor, R., Rahouti, A., Baghouz, A., Li, N., & Cabrera-Guerrero, G. (2020). Towards a Customizable Immersive Virtual Reality Serious Game for Earthquake Emergency Training. *Advanced Engineering Informatics, 46*, 101134.

Co-Authorship Form D

Chapter 6: Teaching Methods in the Customization Framework

The content of this chapter is extracted from:

Feng, Z., González, V.A., Mutch, C., Amor, R., & Cabrera-Guerrero, G. (2020). An Immersive Virtual Reality Serious Game to Train Children in Earthquake Emergency Responses. Submitted to *Journal of Computer Assisted Learning.*

Co-Authorship Form E

Chapter 7: Storytelling Methods in the Customization Framework

The content of this chapter is extracted from:

Feng, Z., González, V.A., Mutch, C., Amor, R., & Cabrera-Guerrero, G. (2020). Exploring Spiral Narratives with Immediate Feedback in Immersive Virtual Rceality Serious Games for Earthquake Emergency Training. Submitted to *Virtual Reality.*

Chapter 1: Introduction

1.1. Introduction

Most people spend a large amount of time living and working in buildings. Buildings serve as a shelter with different functions for people. However, natural hazards such as earthquakes can make buildings dangerous for use with life-threatening environments. In the last three decades, earthquakes have resulted in nearly 0.4 million deaths and one million injuries and affected over 61 million people worldwide (Doocy, Daniels, Packer, Dick, & Kirsch, 2013). It is crucial to have a suitable behavioral response following best practice to maximize the chance of survival and minimize injured people in earthquakes and post-earthquake evacuation (Alexander, 2012; Bernardini, D'Orazio, & Quagliarini, 2016). In New Zealand, national guidelines recommend building occupants to "Drop, Cover and Hold" during earthquakes and to take a list of behavioral responses such as checking and helping others, checking for and mitigating hazards, paying attention to aftershocks (Mahdavifar, Izadkhah, & Heshmati, 2009; New Zealand Ministry of Civil Defence & Emergency Management, 2015; Stuart-Black, 2015). In that respect, educating and training building occupants to be prepared with these behavioral responses is a fundamental process to improve the earthquake resilience and preparedness of communities, which cannot be neglected.

In general, traditional training approaches such as drills have been widely adopted to train building occupants with suitable earthquake safety knowledge (e.g., emergency behavioral responses, self-protection skills, best evacuation practices). However, it has been argued that drills may not effectively prepare trainees (Gwynne et al., 2016). One possible explanation is that trainees are not able to assess their training performance due to the lack of individual feedback after an evacuation drill (Gwynne et al., 2019). Another possible reason is that trainees are not emotionally engaged in the learning processes, which may lead to a reduced effect on attitude and limited change in behaviors (Chittaro, Buttussi, & Zangrando, 2014). In addition, drills also have limited ecological validity. Ecological validity represents the capability of experimental methods, materials, or settings to represent real-world scenarios (Brewer & Crano, 2000). It is almost impossible for drills to reflect earthquake dynamics and effects on buildings and building occupants. It is also unethical to expose trainees to hazardous events. Drills provide a lack of credible training environments for trainees to undertake suitable practice. Therefore, an innovative training approach to overcome the limitations mentioned above is needed.

Recently, Immersive Virtual Reality Serious Games (IVR SGs) have attracted much attention for emergency training (Feng, González, Amor, Lovreglio, & Cabrera-Guerrero, 2018). IVR SGs incorporate two concepts: IVR and SGs. SGs are video games that mainly focus on training, education, or behavioral analysis rather than pure entertainment (Wouters, van der Spek, & van Oostendorp, 2009). SGs have been suggested to deliver effective learning outcomes in different domains, such as professional training, school education, and emergency training (Backlund & Hendrix, 2013; Connolly, Boyle, Boyle, MacArthur, & Hainey, 2012; Girard, Ecalle, & Magnan, 2013). One important feature of SGs is the feedback mechanism, which is lacking in traditional training approaches such as safety cards or drills (Sauvé, Renaud, Kaufman, & Marquis, 2007). With feedback received, trainees can get cues to reflect on their performance, develop perspectives, and establish knowledge (C. I. Johnson & Priest, 2014; Shute, 2008). SGs also facilitate problem-based activities with interactive experience (hands-on practice), which can deepen comprehension and build lasting memories (Greitzer, Kuchar, & Huston, 2007). IVR is a technology which can promote the engaging and immersive capability of SGs. IVR can provide a credible virtual environment that makes users believe they are physically in this virtual world, leading to superior engagement and perception than non-IVR visualizations (LaValle, 2016). In addition, IVR can keep users concentrated on tasks when they are fully immersed in virtual environments, facilitating memory recall (Krokos, Plaisant, & Varshney, 2019). IVR delivers a high-level sense of presence to users by presenting credible environments and simulation (Lovreglio et al., 2018). The close-to-reality simulation allows life-like practice, which can be too dangerous and unethical to be undertaken in the real world. As such, IVR can be integrated into SGs to boost the potentials of SGs, enabling more effective training outcomes.

1.2. Research Problems and Questions

In recent years, various applications of IVR SGs have been studied in the literature, including training, education, and behavioral analysis. However, the study of earthquake emergency training has been insufficiently addressed in the literature (Feng et al., 2018). Unlike fire emergencies, which require people to evacuate buildings as fast as possible, a series of behavioral responses as best practice is recommended in response to earthquakes. The effectiveness of IVR SGs to train people about best practice for earthquakes and post-earthquake evacuation remains unclear.

Beyond that, most of the IVR SG-based training programs focus on ad-hoc training scenarios and single trainee types, delivering identical training outcomes. There is a lack of flexibility for these IVR SG's frameworks to adapt training towards heterogeneous training scenarios and trainee types. In practice, trainees have their own backgrounds, preferences, capabilities, objectives, and learning styles. A fixed training program which fails to consider these characteristics can result in inadequate training outcomes (Kelley, 1969).

As a result, theory and empirical evidence to make a case for a customizable IVR SG framework suited to earthquake emergency training are still lacking. In order to fill these knowledge gaps, the following research questions have been formulated to conduct this research:

Q1: How can an IVR SG suited to earthquake emergency training be developed and implemented?

Q2: How can IVR SG-based earthquake emergency training be customizable suited to different types of trainees and scenarios?

Q3: What are the impacts of a customizable IVR SG framework suited to earthquake emergency training?

1.3. Research Objectives

According to the currently existing problems (see Section 1.2), this research aims to develop a customizable IVR SG suited to earthquake emergency training. This study intends to contribute to the body of knowledge with empirical evidence and provide the theoretical and technical foundations to develop the intended training framework. Therefore, the following research objectives are intended to achieve in this research:

O1: To identify the key factors that contribute to the development and implementation of IVR SGs in the context of emergency training.

O2: To develop a framework including the atomic elements for a customizable IVR SG suited to earthquake emergency training.

O3: To understand the usability aspects and effectiveness related to the IVR SG-based customizable training system in terms of delivering training outcomes, suited to earthquake emergency training.

1.4. Thesis Outline

This thesis consists of ten chapters. This section outlines an overview of each chapter.

Chapter 2 describes the overall research method, covering research stages and strategies.

Chapter 3 presents an extensive literature review, leading to the identification of research questions and the understanding of research topics. A general framework for IVR SG-based emergency studies is proposed.

Chapter 4 validates the proposed general framework through a case study carried out at Auckland City Hospital. This case study is part of a research project funded by the Natural Hazard Research Platform (New Zealand). The author has been of the researchers, engaged in prototype development, data collection and analysis, and leading the writing of the research article presented in this chapter.

Chapter 5 proposes a customization framework for IVR SGs suited to earthquake evacuation training, based on the general framework proposed in Chapter 3 and the adaptive game-based learning suggested in the literature. A customizable IVR SG is developed and tested, validating usability.

Chapter 6 investigates the teaching methods included in the customization framework. A case study assesses the effectiveness of three teaching methods: immediate feedback, post-game assessment, and prior instruction, in terms of increasing knowledge and improving self-efficacy.

Chapter 7 focuses on the storytelling methods in the customization framework. A cases study assesses the impact of two narratives with immediate feedback: linear narratives and spiral narratives, regarding increasing knowledge and improving self-efficacy.

Chapter 8 concludes the research, interprets research contributions, indicates research limitations, and directs future research.

Chapter 2: Research Methodology

2.1. Introduction

This chapter outlines the research methodology applied in this study. In general, the research methodology followed in this research is a constructive research approach, which is "*a research procedure for producing innovative constructions, intended to solve problems faced in the real world and, by that means, to make a contribution to the theory of the discipline in which it is applied*" (Lukka, 2003). This study constructs solutions to solve problems in the real world and contributes to the academic world as well. Typically, the constructive research process includes several phases: find a practically relevant problem which also has research potential, obtain a general and comprehensive understanding of the topic, innovate (i.e., construct) a solution idea, demonstrate that the solution works, show the theoretical connections and the research contribution of the solution concept, and examine the scope of applicability of the solution (Kasanen, Lukka, & Siitonen, 1993). Based on that framework, this study has five major stages, as shown in Figure 2-1. The following sections give an overview of the research stages and research methods applied in each stage. Further details of the applied methods are discussed in the relevant chapters.

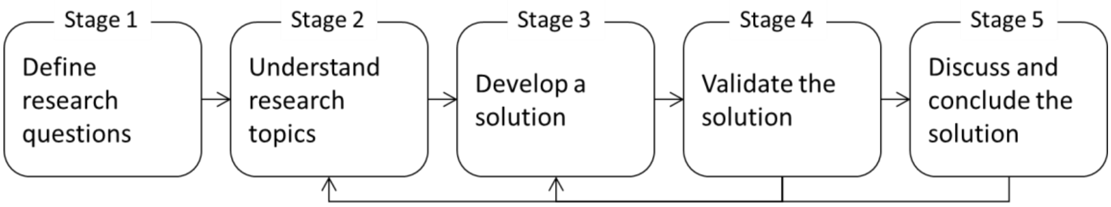


Figure 2-1 Research stages with iteration processes (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007)

2.2. Stage 1: Define research questions

The major research method applied in this stage was a systematic literature review, which has been a method carried out over the entire period of this research. A literature review on the topics of IVR SG-based emergency studies was conducted (see Section 3.1), leading to the identification of research questions and objectives listed in Chapter 1.

2.3. Stage 2: Understand research topics

Stages 2 pays attention to the first research objective (see Section 1.3). The first research method applied in this stage was a systematic literature review. A deep understanding of the IVR SGs-based emergency studies was constructed. Next, a general framework for IVR SG-based emergency studies was developed, connecting the key factors of the development and implementation of IVR SG-based emergency studies. This general framework also laid the foundation for the further development of the IVR SG-based customizable training system for earthquake emergencies. The systematic literature review is reported in Section 3.2. In order to validate the proposed general framework, a case study was carried out, as reported in Chapter 4.

2.4. Stage 3: Develop a solution

Stage 3 focuses on the second research objective (see Section 1.3). Possible factors contributing to a customization framework for IVR SGs suited to earthquake emergency training were investigated to form a customization framework. The general framework proposed in Section 3.2 and the game-based learning (GBL) framework from the literature were the main foundations for the development of the customization framework. The detailed discussion of the customization framework is reported in Chapter 5.

2.5. Stage 4: Validate the solution

At this stage, the third research objective was met (see Section 1.3). Case studies were the main research methods applied at this stage, providing empirical evidence to validate the proposed customization framework. A customizable IVR SG training system targeting earthquake emergencies was developed based on the customization framework. The assessment of the usability of the customizable IVR SG training system is reported in Chapter 5. The assessment of the teaching methods suggested by the customization framework is reported in Chapter 6. The assessment of the storytelling methods suggested by the customization framework is reported in Chapter 7.

2.6. Stage 5: Discuss and conclude the solution

The last stage of this study is to demonstrate the contribution of the customization framework, as well as current limitations and future research directions, as elaborated in Chapter 8.

Chapter 3: Literature Review and a General Framework for IVR SG-based Emergency

The content of this chapter is extracted from:

Feng, Z., González, V.A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., Lovreglio, R., & Cabrera-Guerrero, G. (2020). An Immersive Virtual Reality Serious Game to Enhance Earthquake Behavioral Responses and Post-earthquake Evacuation Preparedness in Buildings. *Advanced Engineering Informatics, 45*, 101118.

Feng, Z., González, V.A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive Virtual Reality Serious Games for Evacuation Training and Research: A Systematic Literature Review. *Computers & Education, 127*, 252-266.

3.1. IVR SGs for emergency training

The use of Serious Games (SGs) for education and training can be traced back to the late twentieth century (Rice, 2007). SGs are identified as video games with serious purposes, such as training, simulation and healthcare, instead of pure entertainment (Michael & Chen, 2006; Susi, Johannesson, & Backlund, 2007). One key reason to use SGs for training is that participants generally feel more engaged and motivated with training processes as compared to other approaches such as watching videos or attending seminars (Papastergiou, 2009). By involving game mechanisms in training, participants are able to interact with objects and environments that assist them to focus on learning content and feedback to enhance learning outcomes (Bellotti, Kapralos, Lee, Moreno-Ger, & Berta, 2013). SGs have been suggested as an effective approach to reinforce traditional training approaches (Gao, González, & Yiu, 2019).

SGs have been applied to various platforms, including mobile devices and desktop computers (Connolly et al., 2012). To provide an advanced immersive and engaging experience, SGs can be integrated with Immersive Virtual Reality (IVR), known as IVR SGs. IVR is a technology that can induce a *"targeted behavior in an organism by using artificial sensory stimulation, while the organism has little or no awareness of the interference"* (LaValle, 2016). IVR can provide a credible virtual environment where participants can explore and behave as close to reality as possible (LaValle, 2016; Sherman & Craig, 2018). Such real-world reactions are essential for behavioral analysis as well as educational applications as participants are expected to shift their behaviors towards recommended ones after training. Krokos et al. (2019) indicated that participants under an IVR condition had better performance in terms of memory recall as compared to non-IVR conditions. Participants were found to be more focused on tasks when they were fully immersed in the virtual environment provided by IVR. Similarly, Chittaro and Buttussi (2015) argued that IVR was beneficial to knowledge retention because of the highly psychological arousal yielded by the high-degree engagement and life-like experience. The synergies that exist between IVR and SG are apparent, which justify the combination of these approaches.

IVR SGs have become a popular tool for emergency training and research. A recent study by Feng et al. (2018) indicated that IVR SGs had been applied to various emergency situations, including fire evacuation (Lin, Zhu, Li, & Becerik-Gerber, 2020; S. Smith & Ericson, 2009), aircraft emergencies (Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015), and earthquakes (Li, Liang, Quigley, Zhao, & Yu, 2017). S. Smith and Ericson (2009) adopted an IVR SG to increase children's motivation towards learning fire safety skills. The findings of this study revealed that participants improved their fire safety knowledge significantly after the training took place. Burigat and Chittaro (2016) applied an IVR SG to train participants about spatial knowledge of an aircraft, in order to undertake an effective evacuation. Participants trained by the IVR SG took less time to evacuate as compared to those trained by safety cards. Li et al. (2017) proposed an IVR SG to train participants in self-protection skills during earthquakes. Participants were asked to detect potential hazards and avoid physical damage during indoor earthquake emergencies. The results suggested that the IVR SG was more effective than other approaches (videos and manuals) in terms of self-protection skills training.

Overall, previous studies have shown that IVR SGs have the potential to generate positive outcomes for emergency training. However, to date, only a few studies have focused on IVR SGs for earthquake behavioral responses and post-earthquake evacuation preparedness, and paid attention to the education dimension to teach and disseminate best evacuation practice (Feng et al., 2018). Little is known about the effectiveness and applicability of IVR SGs to improve the immediate behavioral responses to earthquakes and post-earthquake evacuation preparedness in buildings.

Beyond that, most of the current IVR SGs for emergency training are ad-hoc training solutions that disregard the heterogeneity of trainees and the diversity of scenarios that they can be exposed to. Every trainee is an individual who has his or her own learning style, perceptions, objectives, and competence. Identical training delivered by ad-hoc training solutions has the same content, pedagogical strategies and game mechanisms, but fails to take care of trainees as individuals (Streicher & Smeddinck, 2016).

3.2. A general framework for IVR SG-based emergency studies

3.2.1. Introduction

Today, most people spend a large portion of their time living and working in buildings. However, natural or man-made hazards can make the building environment a dangerous place in which to remain. Proper evacuation responses and behavior during an emergency is a crucial factor to increase survival chance. In general, people are trained and acquire evacuation knowledge (e.g., emergency response, self-protection skills, best practice) through traditional approaches such as videos, posters, seminars, courses, or evacuation drills. However, these traditional approaches may not effectively transmit knowledge (Gwynne et al., 2016). One reason is that after an evacuation drill, building occupants are generally not provided with individual feedback assessing their evacuation behavior (Gwynne et al., 2019). Another reason is that building occupants can be not emotionally engaged in the learning process that may lead to a reduced effect on attitude and limited change in behavior (Chittaro et al., 2014). In fact, Yang et al. (2011) pointed out that real-life evacuation behavior is different from experiments such as evacuation drills, which means that current evacuation models still have limitations when they are the basis of evacuation training and research. Evacuation drills also have other limitations such as being costly in time and resources by interrupting building occupants' routines, and being not able to present hazards (Gwynne et al., 2016; Gwynne et al., 2019; Silva, Almeida, Rossetti, & Coelho, 2013). Therefore, there is a need to investigate innovative and more effective approaches to overcome the limitations mentioned above (Kobes, Helsloot, de Vries, & Post, 2010). Such innovations should be able to transmit evacuation knowledge of building occupants towards one more effective and efficient.

In that regard, Serious Games (SGs) have attracted much attention to pedagogical research recently (Connolly et al., 2012). SGs are video games whose primary purposes are training and education, not entertainment per se (Wouters et al., 2009). It is argued that by playing SGs, participants can gain and retain knowledge more effectively than by using traditional learning methods (Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013). SGs can enhance the capability of traditional approaches to delivering evacuation knowledge (I. Mayer, Wolff, & Wenzler, 2013; Zhou, Chang, Pan, & Whittinghill, 2016). In turn, there is another technology that can promote the engaging capabilities of SGs, namely Immersive Virtual Reality (IVR). IVR allows participants to be fully immersed in virtual environments that can provide greater engagement and perception than videos, text-based papers or 2D games (Gao, Gonzalez, & Yiu, 2017; Lovreglio et al., 2017). The combination of IVR and SGs encourages participants to retain knowledge longer than traditional approaches due to the fact that they benefit from full engagement, and high emotional and physiological arousal (Chittaro & Buttussi, 2015). As such, an increasing number of studies have investigated the combination of IVR and SGs for evacuation training and behavior assessment. A systematic understanding of how IVR SGs have been developed and implemented for evacuation training and research is necessary.

This paper introduces a systematic literature review regarding IVR SGs oriented toward building evacuation processes tailored to indoor emergencies. As a result, a conceptual framework to guide the development and implementation of such IVR SGs is proposed. Thus, the key factors contributing to successful and comprehensive development and implementation of IVR SGs tailored to building evacuation are identified and connected.

3.2.2. Background

3.2.2.1. Serious Games

SGs have become a popular training and behavior analysis tool in the last decades (Connolly et al., 2012). The term "serious game" usually represents a video game whose primary purpose is education, training, simulation, socializing, exploring, analyzing and advertising, rather than pure entertainment (Michael & Chen, 2006). Susi et al. (2007) suggested that SGs represent "*the application of gaming technology, process, and design to the solution of problems faced by businesses and other organizations. SGs promote the transfer and cross-fertilization of game development knowledge and techniques in traditionally non-game markets such as training, product design, sales, marketing, etc.*" Note that SGs are not only able to transfer knowledge, but are also able to fulfil other objectives such as behavior analysis (Issa & Zhang, 2015) and rehabilitation healthcare (Schonauer, Pintaric, Kaufmann, Jansen-Kosterink, & Vollenbroek-Hutten, 2011).

One of the primary objectives of SGs is to educate participants (Connolly et al., 2012), and SGs have been investigated widely in different domains for education purposes. Connolly et al. (2012) suggested that SGs can be applied to acquiring knowledge, understanding content, or developing specific skills. For instance, Johnson and Wu (2008) explored the capability of SGs to facilitate teaching foreign languages. Muratet et al. (2009) proposed an SG to improve programming skills, and Sliney and Murphy (2008) and Diehl et al. (2011) developed SG prototypes for medical training. There has also been a number of studies focused on emergency training. SGs have been implemented for emergency training in the oil industry (I. Mayer et al., 2013; Metello, Casanova, & Carvalho, 2008), terrorist attacks (Chittaro & Sioni, 2015), fire evacuation (Chittaro & Ranon, 2009; Sacfung, Sookhanaphibarn, & Choensawat, 2014), disaster evacuation (D. Cohen et al., 2012), and earthquake evacuation (Barreto et al., 2014; Tanes & Cho, 2013). All these studies suggested that SGs are a promising tool for education and training purposes. One explanation is that participants can recall more effectively what they have learned compared to traditional learning approaches (Bartolome, Zorrilla, & Zapirain, 2011). Papastergiou (2009) argued that SGs could positively motivate participants during the learning experience. Another explanation is that participants have the chance to interact with the environment and get immediate feedback from the SGs to rectify any incorrect responses and to strengthen knowledge (Lovreglio et al., 2018). SGs are an effective tool to support and enhance traditional training tools (Gao et al., 2019).

Another objective of SGs is to investigate human behavior (Connolly et al., 2012). Their gaming structure enables tracking and recording participants' decisions and behavior during a game experience. By collecting and analyzing behavioral data, it is possible to reveal behavioral motivation, validate behavioral models, explore decision-making, recognize behavioral patterns, and assess the responses under various controlled conditions (Connolly et al., 2012). For instance, one study (Chittaro & Ranon, 2009) adopted VU-Flow tool (VU-Flow provides a set of interactive visualizations that highlight interesting navigation behavior of single or groups of moving entities that were in the virtual environment together or separately) with the game to track and visualize participants' fire evacuation routes, in order to understand the evacuation navigation patterns. In another study (Li et al., 2017), participants' awareness of potential hazards during an earthquake was simulated by tracking visual attention to falling and fragile objects inside the game environment. Therefore, SGs have potential to allow the understanding of behavioral patterns and behavior changes beyond educational and training aspects.

3.2.2.2. Immersive Virtual Reality

Virtual Reality (VR) technology has rapidly evolved in recent years, bringing a wide range of application areas due to its flexibility to adapt to different problems and domains (LaValle, 2016). This has also brought different interpretations of what VR is. In this paper, we refer to VR to that experience in which participants are fully immersed into the virtual environment provided by head-mounted displays (HMD) or projection-based displays (PBD) (Sharples, Cobb, Moody, & Wilson, 2008). The most popular PBD application is the CAVE system (cave automatic virtual environment), which displays images on screens formed up a 3D immersive room, providing an immersive visual and interaction experience (Cruz-Neira, Sandin, & DeFanti, 1993). Figure 3-1(a) shows the HMD system and Figure 3-1(b) shows the PBD system.

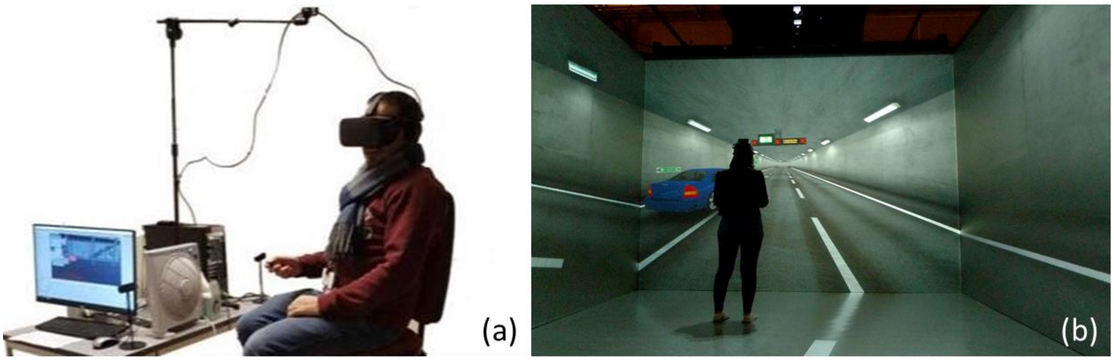


Figure 3-1 (a) HMD system (Lovreglio et al., 2018) ; (b) PBD system (Ronchi et al., 2016)

We use the term “immersive virtual reality (IVR)” to distinguish it from other forms of VR. IVR can be defined as “*Inducing targeted behavior in an organism by using artificial sensory stimulation, while the organism has little or no awareness of the interference*” (LaValle, 2016). According to this definition, once participants are immersed in the virtual environment, they can feel they are physically inside this environment, even though such environment is artificially simulated. This virtual environment may become very realistic, making it very difficult for individuals to differentiate between the virtual and the real world (LaValle, 2016). Therefore, IVR has the potential to allow participants to behave and react as close as possible to reality (Sherman & Craig, 2018).

IVR can create a number of benefits resulting from its applications. IVR has been demonstrably effective in communicating the cultural content of a museum exhibition (Carrozzino & Bergamasco, 2010). It also can perform social interaction through behavior and context to improve the learning environment (Bailenson et al., 2008). In another study (Blascovich et al., 2002), IVR was proposed as a social psychological research tool to lessen the trade-off between mundane realism and experimental control. IVR is also an ideal tool for exploration, training, and education (Psotka, 1995). As a result, its potential for emergency training and education has been investigated (Li et al., 2017; Sharma, Jerripothula, Mackey, & Soumare, 2014). Shendarkar et al. (2008) suggested that it is more effective to conduct emergency management by using IVR due to its capability to model human behavior with a high degree of fidelity. S. Smith and Ericson (2009) revealed that IVR could enhance the enthusiasm of children for fire-safety skills training by improving their engagement with the learning environment.

A recent study by Krokos et al. (2019) found that IVR can provide better memory recall ability compared to non-IVR conditions. Their research showed that participants felt more focused on the task resulting from better immersion experience. In addition, the majority of the participants also claimed that the sense of the spatial awareness enhanced by IVR was critical to their success. Given that, IVR experience can influence SGs elements to make the learning and behavior outcome significant resulting from its special cognitive process. SGs can be enhanced for training and education purposes using IVR principles, and thus, they can be regarded as IVR SGs. The involvement of IVR principles in SGs provides a higher degree of engagement when compared to non-IVR SGs (Gao et al., 2017; Lovreglio et al., 2017). Chittaro and Buttussi (2015) suggested that IVR SGs can improve knowledge retention and psychological arousal in aviation safety training. One possible interpretation is that IVR SG frameworks have the capability to generate “realistic scenarios” and hazards, which in turn create more "realistic feelings" in participants. The effect is to make evacuation simulation highly engaging so that participants can get better hazards perception and risk awareness. Apart from the benefit of training and education, IVR SGs are also valuable tools for analyzing human behavior during different emergencies such as fire (Kinateder, Ronchi, Nilsson et al., 2014) or earthquakes (Li et al., 2017). Kinateder et al. (2014) argued that IVR SGs allow a safe study of occupant behavior in scenarios that would be too dangerous to implement in the real world evacuation drills such as dense smoke or falling objects.

However, IVR SGs have limitations. One significant issue is that a VR environment may induce motion sickness (Hettinger & Riccio, 1992). Sharples et al. (2008) argued that there is a high chance of producing virtual reality induced side effects (motion sickness) by using head-mounted displays. A possible explanation for this side effect is that participants suffer from sensory conflicts when they view compelling visual representations of self-motion in the IVR environment with physically stationary self-body (Hettinger & Riccio, 1992). Regan and Price (1994) stated that 61% of participants felt symptoms of malaise, such as dizziness, nausea, and headache during the simulation. However, Lovreglio et al. (2018) reported that only 5% of participants felt motion sickness in their study, that was benefited from the high quality of navigation systems. Therefore, the thoughtful design of IVR SGs is necessary to minimize these side effects.

Various applications of IVR have been explored by researchers. However, so far, there has been no literature review systematically assessing the combination of IVR and SGs for building evacuation purposes. Given that, there is still a need to understand how to systematically develop and implement IVR SGs for building evacuation training and research.

3.2.3. Research Design

Inspired by the capabilities and potential of IVR, this study is focusing on the combination of IVR and SGs. This study aims to provide insight into the development and implementation criteria of IVR SGs oriented towards indoor evacuation processes. A conceptual framework is expected to be generated as the main outcome of this study.

In order to achieve this objective, a systematic literature review was conducted to comprehensively explore the existing IVR SGs tailored to building evacuation training and research. Qualitative data analysis was carried out to identify the empirical evidence on the essential factors contributing to effective and robust development and implementation of IVR SGs suited to indoor evacuation training and research. As a result, a framework was generated, integrating the analyzed evidence providing guidelines for future research.

3.2.4. Systematic Literature Review

The systematic literature review was conducted in accordance with the framework recommended by Khan et al. (2003) and Thomé et al. (2016). This review included five stages: formulating the research problems, identifying relevant work, assessing the quality of studies, summarizing the evidence, and interpreting the findings.

3.2.4.1. Formulating the Research Problems

Rüppel and Schatz (2011) elaborated a triadic game design approach to SGs that included three interdependent worlds that need to be balanced during the design process: reality (how the game is connected to the physical world), meaning (what value needs to be achieved), and play (how to create playful activities). We investigated these three major aspects in this study. The investigation allowed us to answer two main research questions. In order to get a detailed understanding of these two main questions, eleven sub-questions were formulated. Table 3-1 shows the question, sub-questions and assessed aspects.

Table 3-1 Systematic literature review research questions

Main research questions	Sub-questions	Assessed aspects
MQ1: What are the outcomes and measures for implementing IVR SGs in evacuation study?	SQ1: What are the learning outcomes? SQ2: What are the learning measures?	Meaning: Pedagogical impact
	SQ3: What are the behavior outcomes? SQ4: What are the behavior measures?	Meaning: Behavioral impact
	SQ5: How participation experience can be evaluated?	Play: Participation experience
MQ2: What are the essential elements for developing IVR SGs in evacuation study?	SQ6: What are the teaching methods? SQ7: What are the navigation solutions? SQ8: What are the sensations stimulated? SQ9: What are the narrative methods to encourage the participants to follow the game storyline, and complete it?	Play: Hardware and software system

	SQ10: Are there non-player characters (NPCs) and how do they contribute?	
	SQ11: What are the hazards simulated?	Reality: Software system

In general, two types of impacts of IVR SGs for evacuation training and research can be summarized; one is a pedagogical impact, and the other is a behavioral impact (Lovreglio et al., 2018). Pedagogical impacts refer to the effects resulting from education or training through the use of IVR SGs, manifested in acquiring knowledge or improving skills for participants. Behavioral impacts mean that human behavior is investigated by using IVR SGs due to the fact that they can be used as effective behavioral analytical tools. Accordingly, participants are exposed to challenges, tasks and specific situations in the IVR SG environment prompting their response in the form of actions or decisions. SQ1 and SQ2 were formulated to explore the pedagogical impact, while SQ3 and SQ4 were for the behavioral impact. Apart from that, participation experience is also an important aspect that can be used to support and refine the prototype (Chittaro & Buttussi, 2015). On that basis, SQ5 was formulated to explore how to measure the participation experience. In terms of the gaming environment development, it includes various components, which could be either the hardware system and software system (Rüppel & Schatz, 2011). Therefore, RQ6 to RQ11 were formulated to discover the specific details of different systems, including the teaching methods, navigation solutions, sensory stimulation, narrative methods, NPCs, and hazards simulation.

3.2.4.2. Identifying the Relevant Work

The eligible papers need to include three major concepts, namely, immersive virtual reality, serious games, and evacuation training and research. IVR is a mechanism added onto SGs, while SGs are still the key content in the eligible papers which cover existing knowledge on SGs to achieve their primary research aims.

Eligible papers included in the systematic literature review were collected from journals and conference proceedings. The papers were recovered from the following databases: Scopus and Engineering Village. Scopus is the largest abstract and citation database of peer-reviewed literature, covering 12,464 titles on social sciences and 13,312 titles on physical sciences (Scopus, 2020). Engineering Village is an index of the most comprehensive engineering literature, including journals, standards, conference proceedings, dissertations, books, and patents (Engineering Village, 2020). Meanwhile, another approach called snowballing (retrieve relevant papers based on target papers' references list or paper citing) (Wohlin, 2014) was also adopted with Google Scholar, which indexes most scholarly literature, as a complementary method to cover any missing papers.

There is an inconsistency in the terminology used in the literature. For instance, virtual reality, virtual environment, virtual simulation, or VR can all represent the content of immersive virtual reality. To get the maximum coverage of publications, we conducted searches using the following search string: "virtual reality" (enclose the phrase in braces or quotes to find papers that contain the exact phrase) OR "virtual environment" OR "virtual simulation" OR VR AND evacuation. For Scopus, the search fields were article titles, abstracts, and keywords. For Engineering Village, the search fields were subjects, titles, and abstracts. The searches were not limited to any other constraints, such as language or time span. The searches were conducted on 22 January 2018 and yielded a total of 567 results (including duplicates), 233 of which were from Scopus (www.scopus.com) and 334 from Engineering Village (www.engineeringvillage.com).

After duplicates were removed, a filtering process was carried out following a framework called Preferred Reporting Items for Systematic Literature Reviews and Meta-Analyses (PRISMA) (Moher, Liberati, Tetzlaff, & Altman, 2009). The papers were filtered in accordance with the following inclusion and exclusion criteria: firstly, papers were excluded if there were no IVR SG-based evacuation training and research related to terms in the titles or abstracts. Subsequently, the rest of the papers' full texts were assessed for eligibility. The eligible papers were included if they met all the following criteria:

(i) An IVR SG prototype was proposed, or an existing IVR SG prototype was evaluated and analyzed;

(ii) An experiment was carried out to gather the learning or behavior outcome;

(iii) The data analysis of the outcome was carried out to evaluate and validate the prototype.

The papers were excluded if meeting one of the following criteria:

(i) There was no immersive virtual reality principle involved in the prototype because there are other forms of VR (e.g., flat screens showing virtual environment) which do not involve full participant immersion;

(ii) Only the theories, concepts, frameworks, or proposals were discussed without following up experiments or case studies. This study aims to investigate IVR SGs for evacuation from development to implementation. The implementation stage is an important step not only to evaluate and validate IVR SGs but also to implement them into practice and influence a large number of evacuees.

After the filtering process, snowballing was adopted to identify additional papers. The snowballing was based on the previous filtered results after PRISMA. Both backward and forward snowballing was carried out on Google Scholar. The inclusion and exclusion criteria were the same as the ones adopted in the PRISMA framework. Figure 3-2 shows diagrammatically the above-mentioned methodological process derived from the PRISMA framework.

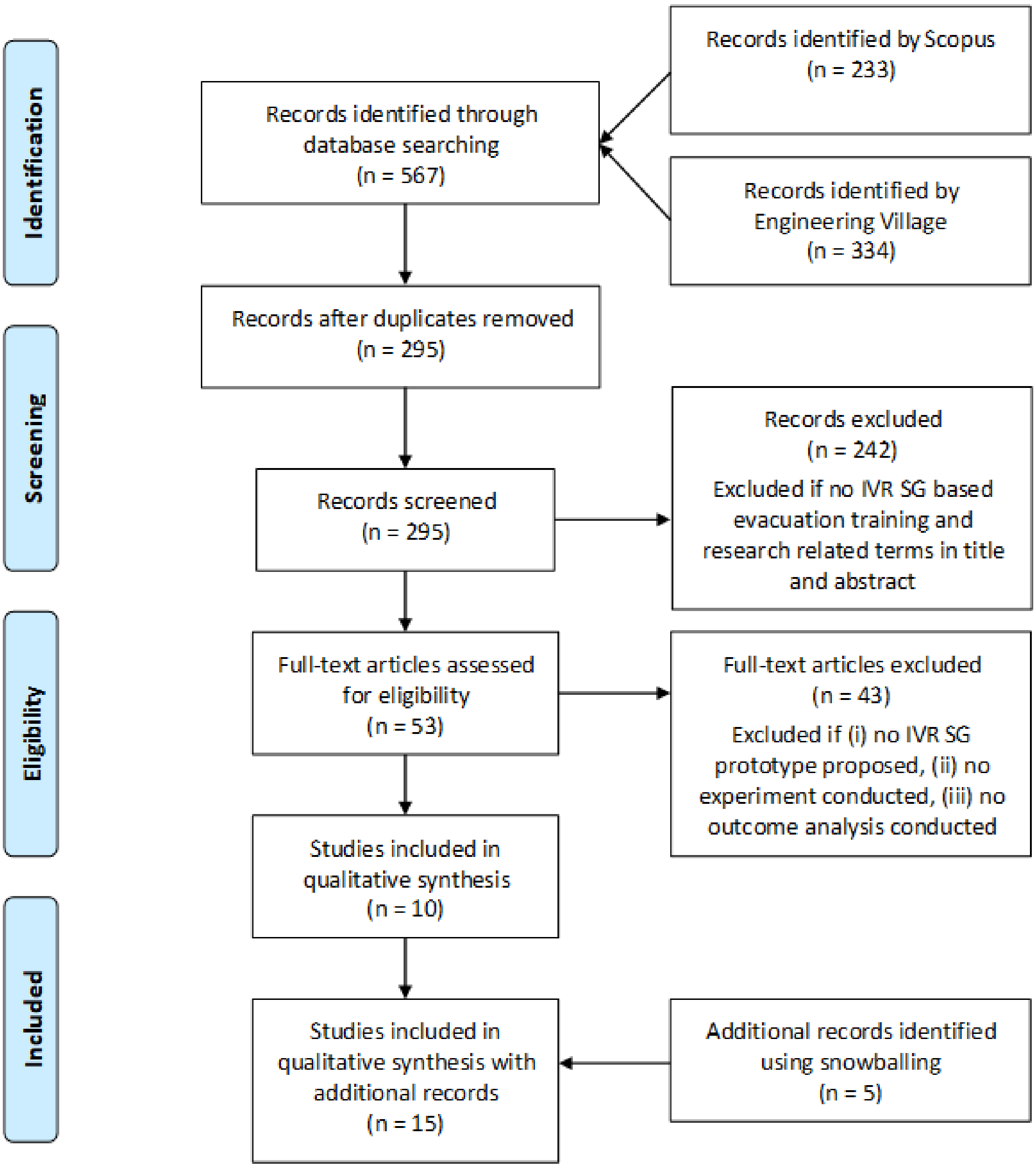


Figure 3-2 Study selection process

As a result, 15 papers were identified as being relevant to this systematic literature review, which are (Andrée, Nilsson, & Eriksson, 2016; Aoki, Oman, & Natapoff, 2007; Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015; Cosma, Ronchi, & Nilsson, 2016; Duarte, Rebelo, Teles, & Wogalter, 2014; Kinateder, Ronchi, Gromer et al., 2014b; Kinateder, Müller, Jost, Mühlberger, & Pauli, 2014a; Kinateder et al., 2015; Li et al., 2017; Meng & Zhang, 2014; Ronchi et al., 2015; Ronchi et al., 2016; S. Smith & Ericson, 2009; Zou, Li, & Cao, 2016). Among these papers, 14 were published in journals, with the remaining one published as a conference proceeding. Most of the papers were published after 2014, with only two were published in 2007 and 2009. One possible reason may be that IVR technology has become more popular only in recent years, although the concept of IVR itself can be traced back to the 1980s (Fisher, McGreevy, Humphries, & Robinett, 1987). Another interesting finding is that most of the IVR SGs for evacuation training and research have been carried out in Europe, with only three conducted in China and another two in the US.

3.2.4.3. Assessing the Quality of Studies

After the eligible papers were identified, a scoring process was conducted to assess the quality of the papers. The scoring criteria were derived from the quality assessment approach adopted by Connolly et al. (2012). To be specific, each paper was scored by two authors based on the three assessment questions asked below:

(i) How appropriate is the prototype design for addressing the questions of this review?

(ii) How appropriate are the methods and analysis for addressing the questions of this review?

(iii) How relevant is the focus of the study for addressing the questions of this review?

Each dimension score ranged from 1 to 3 where 1 meant low quality, 2 meant medium quality, and 3 meant high quality. Each of the three dimensions' scores was summed to get a total score for each paper. As a result, each paper received two total scores from two raters, and the mean for these two scores was calculated to get the final score of each paper. Possible final scores ranged from 3 to 9, where 3 stood for low quality, and 9 stood for high quality.

Each of the 15 papers was coded, and Figure 3-3 shows a histogram of the final scores. The mean for the 15 papers' ratings is 6.63, and the mode is 6. Twelve papers rated 6 or over were considered as higher quality papers that provided stronger empirical evidence regarding this review's objective. All the papers are summarized in Appendix A with high-quality papers being highlighted, showing the names of the authors, years published, objectives of the study, methods, results and conclusions.

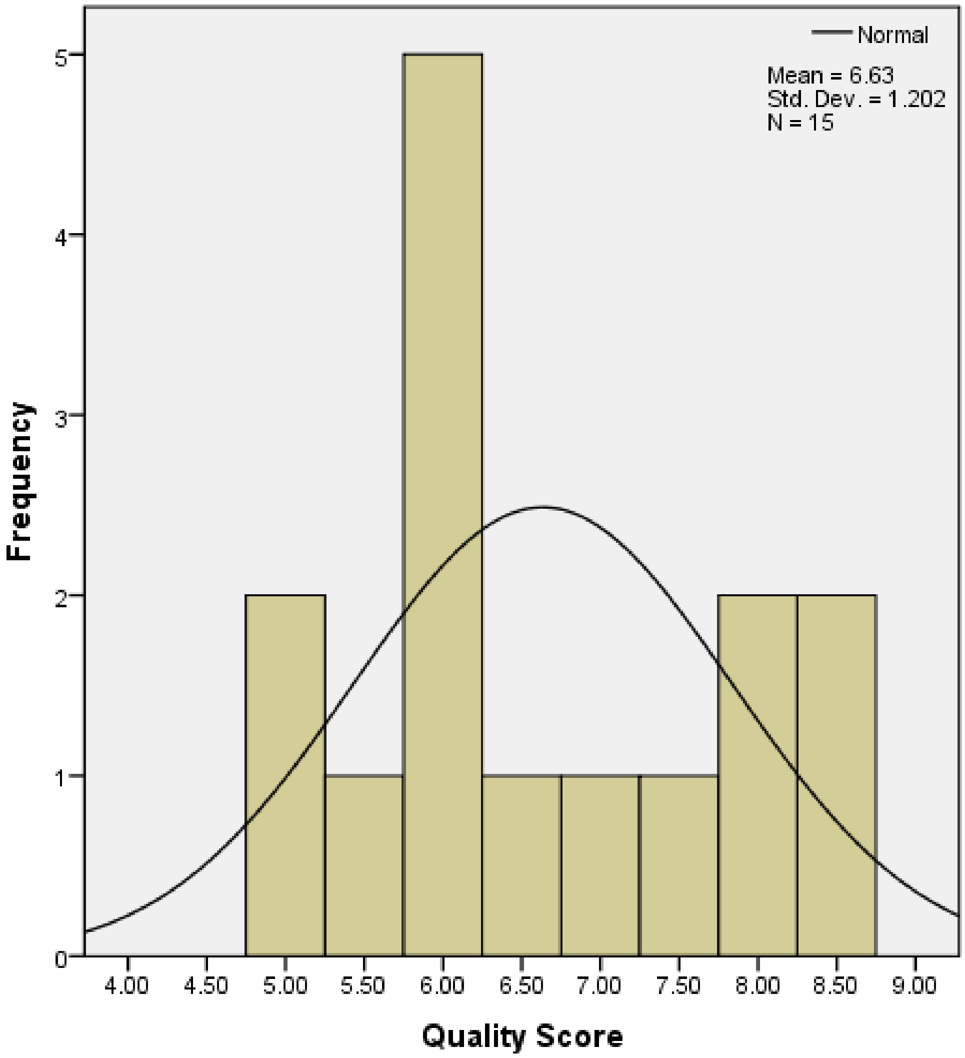


Figure 3-3 Quality scores for eligible papers

3.2.4.4. Summarizing the Evidence

The eligible papers were coded and analyzed using a data extraction spreadsheet that included the research aspects and questions mentioned above. Given that there is only a small number of eligible papers, all the 15 papers are discussed in-depth in this section.

3.2.4.4.1. Gaming Outcomes and Measures

Two types of outcomes were identified in Section 3.2.4.1, namely, pedagogical outcomes and behavioral outcomes. Table 3-2 shows the number of papers that addressed the different outcomes in terms of the simulated emergencies of the games.

Table 3-2 Gaming outcomes and impacts

Simulated Events	Pedagogical Outcomes	Behavioral Outcomes	Pedagogical and Behavioral Outcomes
Tunnel Fire		(Cosma et al., 2016; Kinateder et al., 2014a; Kinateder et al., 2014b; Kinateder et al., 2015; Ronchi et al., 2015; Ronchi et al., 2016)	
Building Fire	(S. Smith & Ericson, 2009)	(Andrée et al., 2016; Duarte et	

		al., 2014; Meng & Zhang, 2014)	
Aviation Emergency	(Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015)		
Spacecraft Emergency	(Aoki et al., 2007)		
Building Earthquake	(Li et al., 2017)	(Li et al., 2017)	(Li et al., 2017)

Among the eligible papers, there were four papers that were identified to use IVR SGs as a pedagogical tool, while nine papers were identified to use IVR SGs as a behavioral tool. One paper used IVR SGs as both a pedagogical and behavioral tool.

3.2.4.4.1.1. Pedagogical Impact

In five studies, IVR SGs were implemented as pedagogical tools (Aoki et al., 2007; Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015; Li et al., 2017; S. Smith & Ericson, 2009).

In the reviewed papers, various evacuation knowledge was delivered as learning outcomes. In total, three types of knowledge were identified by this review, namely evacuation best practice (Chittaro & Buttussi, 2015; S. Smith & Ericson, 2009), self-protection skills (Li et al., 2017), and spatial knowledge (Aoki et al., 2007; Burigat & Chittaro, 2016).

S. Smith and Ericson (2009) encouraged participants to identify potential fire hazards. Following that, participants were educated on the best practice for evacuation in case of fire at home. A significant increase in measured fire-safety knowledge after training was observed in the results. Chittaro and Buttussi (2015) provided the entire evacuation protocol for an aviation emergency ranging from turbulence instability response to evacuation into a life raft after a forced landing at sea, targeting passengers. As a result, participants were trained with a total of ten learning items associated with the best practice for an aviation emergency. Results showed that IVR SG had better performance in terms of knowledge retention after one week compared to the safety card method. This is critical for people to recall knowledge of proper evacuation behaviors in real emergencies.

Apart from the studies focusing on multiple knowledge (e.g., best practice), three other papers focused on single knowledge transfer. Li et al. (2017) proposed using IVR SG as a training tool to teach participants how to protect themselves in common indoor environments during an earthquake. On average, participants trained by IVR SG performed better than those trained by safety videos or manuals in terms of hazard awareness and avoidance. Burigat and Chittaro's (2016) main learning goal was to let participants acquire spatial knowledge to permit effective evacuation of an airplane. Results showed that participants trained by IVR SG obtained better spatial knowledge than those trained by safety cards. Aoki et al. (2007) investigated spatial skills influenced by relative body orientation during IVR training in a spacecraft evacuation. Results showed that local training (visually upright relative to the "local" module) enabled landmark and route learning, while station training (constant orientation irrespective of local visual vertical) improved sense of direction and performance in low visibility.

The reviewed papers show that IVR SGs have the capability of delivering a certain amount of evacuation knowledge as learning outcomes, even as complex as best practice with multiple learning items. The expected learning outcomes, which are the basis for game storyline and narrative, should be defined prior to IVR SGs development.

After the IVR SGs training session, the learning performance of all participants mentioned previously (Aoki et al., 2007; Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015; Li et al., 2017; S. Smith & Ericson, 2009) was assessed bearing in mind the IVR SGs learning outcomes. Several assessment measures were recognized, and these are shown in Table 3-3.

Table 3-3 Learning measures of IVR SGs

Learning Measures	Knowledge Acquisition	Knowledge Retention	Both Knowledge Acquisition and Retention
Questionnaire	(S. Smith & Ericson, 2009)		
Open-ended question interview	(Aoki et al., 2007; Chittaro & Buttussi, 2015)	(Chittaro & Buttussi, 2015)	(Chittaro & Buttussi, 2015)
Paper-based test	(Burigat & Chittaro, 2016)		
Logged game data (e.g., evacuation time, damages received)	(Aoki et al., 2007; Burigat & Chittaro, 2016; Li et al., 2017)	(Li et al., 2017)	(Li et al., 2017)

In order to assess participants' knowledge acquisition, S. Smith and Ericson (2009) measured how many correct answers were given to the questions related to evacuation safety knowledge before and immediately after the IVR SGs training. Aoki et al. (2007) asked participants to verbally describe the configuration of the spacecraft after the training, thus, to assess their spatial knowledge. Chittaro and Buttussi (2015) asked participants to orally answer the open-ended questions related to evacuation safety knowledge before and immediately after the IVR SGs training to avoid suggesting possible answers such as a multiple-choice questionnaire would do. In another study, Burigat and Chittaro (2016) adopted a paper-based approach using maps to let participants mark positions of exits and their initial seats as they aimed to gain spatial knowledge. In addition to the paper-based approach, Burigat and Chittaro (2016) also recorded the evacuation time of participants to evaluate their learning results. Aoki et al. (2007) recorded evacuation time, numbers of turns and errors to assess participants' spatial knowledge acquisition. Li et al. (2017) used logged game data in terms of the physical damage received during the training session to evaluate participants' learning performance, thus, to provide feedback in order to improve self-protection skills. In accordance with learning outcomes, different assessment measures can be applied.

S. Smith and Ericson (2009) argued that there was a significant improvement in measured fire-safety knowledge after training. Chittaro and Buttussi (2015) revealed that IVR SGs are superior to traditional approaches regarding knowledge retention. This is the fundamental requirement for survival during evacuation because people need to retain correct evacuation procedures over a long-time span so that to apply them whenever facing an emergency. Li et al. (2017) and Burigat and Chittaro (2016) found IVR SGs could produce better knowledge transfer than traditional approaches.

Therefore, it can be concluded that IVR SGs are effective in delivering considerable evacuation knowledge, no matter whether it is multiple knowledge (e.g., best practice) or single knowledge (e.g., spatial skill).

After running the entire IVR SGs training session, researchers from two studies (Chittaro & Buttussi, 2015; Li et al., 2017) conducted second-round tests one week later to assess participants' knowledge retention. With the learning measures adopted for assessing knowledge acquisition in the first tests, the same measures were applied again in the second tests to observe the changes in evacuation knowledge between the first and second tests so that the knowledge retention can be evaluated. Both studies showed that in the context of evacuation training, participants that received training by IVR SGs had better performance in terms of knowledge retention compared to those that were trained using traditional approaches.

3.2.4.4.1.2. Behavioral Impact

In ten studies, the IVR SGs approach was adopted as a behavioral analysis tool (Andrée et al., 2016; Cosma et al., 2016; Duarte et al., 2014; Kinateder et al., 2014a; Kinateder et al., 2014b; Kinateder et al., 2015; Li et al., 2017; Meng & Zhang, 2014; Ronchi et al., 2015; Ronchi et al., 2016). In terms of the behavioral outcomes and measures, they are rather diverse due to the various purposes of behavioral analysis. Table 3-4 shows the different purposes recognized by this review.

Table 3-4 Behavioral outcomes of IVR SGs

Behavioral Outcomes		Description
Evacuation facility validation	(Andrée et al., 2016; Cosma et al., 2016; Ronchi et al., 2016)	Test and validate different evacuation facility designs and installations

Behavioral compliance	(Duarte et al., 2014)	Investigate whether participants follow the evacuation instructions
Hazard awareness	(Li et al., 2017)	Investigate whether participants can notice hazards in the environment
Behavior validation	(Ronchi et al., 2015)	Validate a hypothetical behavioral model
Social influence	(Kinateder et al., 2014b; Kinateder et al., 2014a)	Investigate social influence on evacuation behavior
Behavior recognition	(Kinateder et al., 2015)	Recognize different behavior under different evacuation conditions
Way-finding behavior	(Meng & Zhang, 2014)	Explore evacuation way-finding behavior

Cosma et al. (2016) investigated the impact of evacuation lighting systems for rail tunnel evacuation. Participants were exposed to an emergency evacuation scenario, and their movement paths and evacuation time were recorded to evaluate different way-finding lighting installations impacts. Results showed that both dynamic alternate and continuous way-finding lighting systems had a positive impact on evacuation, and no significant differences were found between these two systems. Ronchi et al. (2016) put participants in IVR SGs to test different designs of flashing lights at emergency exit portals. Then participants were asked to finish a questionnaire to provide recommendations on different portal designs. Various variables were investigated. Results suggested that participants preferred green and white flashing lights rather than blue lights, at a flashing rate of 1 and 4 Hz rather than 0.25 Hz, a light-emitting diode light source rather than single and double strobe lights. No significant preference difference was found between different numbers and layouts of lighting on portals. Andree et al. (2016) proposed using IVR SGs to validate the evacuation procedures for elevators in high-rise buildings by analyzing participants' exit choices and waiting times. Evacuation elevators are deployed to tackle the issues of fatigue and movement constraints (e.g., using a wheelchair or crutch) during the vertical evacuation in high-rise buildings (Andrée et al., 2016). Results suggested that participants would more likely evacuate by elevators influenced by the green flashing way-finding lighting system. Moreover, participants tended to wait for elevators either a limited time (<5 min) or a long time (>20min). Duarte et al. (2014) studied participants' behavioral compliance by counting the number of times participants followed the directions indicated by the evacuation signs. Results showed that evacuation signs were effective in changing evacuees' behavior. Li et al. (2017) investigated whether participants noticed hazards around them by tracking their visual attention. The percentage of dangerous objects noticed by participants during the IVR SGs was calculated to assess participant behavior in response to an emergency. Results indicated that the participants trained by IVR SG were able to notice more hazards than those trained by other approaches such as videos and manuals. Ronchi et al. (2015) proposed an IVR SG to perform evacuation model validation by comparing participants' actual movement paths to the hypothetical paths. Results showed that hypothetical paths based on the shortest distance employed by evacuation models might be over-simplified compared to actual movement paths. Kinateder et al. (2014b; 2014a) used IVR SGs to investigate social influences on evacuation behavior. The non-player characters (NPCs) inside the IVR SGs undertook social interactions with participants. Participants' behavior was analyzed by evaluating destination choice, movement pathway, and evacuation time to explore social influences on their decision-making processes. Results indicated that other evacuees had a strong influence on participants evacuation behavior. Participants were more likely to react similarly to the NPCs. Kinateder et al. (2015) let participants face burning dangerous goods vehicle or burning heavy goods vehicle to explore different outcomes of evacuation behavior under different conditions. Results showed that participants had similar patterns of evacuation behavior under both conditions, while they still perceived significantly more danger in the burning dangerous goods condition. Meng and Zhang (2014) argued that IVR SGs were capable of exploring the evacuation way-finding behavior by analyzing participants' evacuation time, route choice, and movement pathway. Results suggested that participants had poor way-finding performance in a fire emergency, which may be resulted from high physiological and psychological stress.

Benefit from the nature of IVR SGs, absorbing gaming environment is able to induce participants to react as close as to the reactions in real life. The empirical evidence from the reviewed papers demonstrates that IVR SGs are a promising tool to analyze participants' behavior for different purposes. As for the behavior measures, the choose of different approaches is heavily relying on different expected behavior outcomes. Considering that, IVR SGs need to integrate the corresponding functions to carry out the analysis for pre-defined behavior outcomes.

3.2.4.4.1.3. Participation Experience

A few studies investigated participation experience. It can be categorized into two different aspects, which are self-reported psychological assessment and device-based physiological assessment. In terms of the psychological aspect, Chittaro and Buttussi

(2015), S. Smith and Ericson (2009), Burigat and Chittaro (2016), and Meng and Zhang (2014) introduced questionnaires to assess self-reported fear, engagement, stress, usability and workload. Zou et al. (2016) applied the positive affect and negative affect scale (PANAS) to let participants describe emotions. In relation to the physiological aspect, Chittaro and Buttussi (2015) provided an electrodermal activity sensor (EDA) to track skin conductance levels (SCL) to evaluate fear and anxiety (an increase in SCL indicates arousal), and a photoplethysmograph sensor (PPG) to obtain blood volume pulse amplitude (BVPA), which can be employed as an index of sympathetic arousal (a decrease in BVPA indicates increased arousal). Meng and Zhang (2014) tracked participants skin conductivity and heart rate in order to give an insight into the stress of participants when they were undertaking an IVR SGs session. Zou et al. (2016) applied a multichannel physiological recorder to track participants skin conductivity in order to measure emotional responses of participants when facing a fire emergency.

Participation experience can be measured to evaluate IVR SGs. Self-reported psychological assessment can be applied to investigate the usability of IVR SGs while device-based physiological assessment can be applied to reveal the subconscious activities as strong feasibility evidence to support IVR SGs.

3.2.4.4.2. Gaming Environment

Six major components for the gaming environment were raised in the previous questions, namely teaching method, navigation solution, narrative method, hazard simulation, sensory stimulation, and non-player characters (NPCs).

3.2.4.4.2.1. Teaching Methods

Two types of teaching methods were used to deliver evacuation knowledge. One was to give feedback after the response to an event (Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015; Li et al., 2017). In this way, participants could know if their behavior was right or wrong after making a decision to an event; thus, participants could learn from their mistakes. Another one was to provide instructions before the response to an event (Aoki et al., 2007; S. Smith & Ericson, 2009). In other words, participants were told what to do before dealing with an event, thus, to learn the appropriate behavior accordingly.

Regarding feedback after the response, there are two forms of feedback identified. One is immediate feedback. To be specific, if participants make an inappropriate action, they will be informed by the written or verbal messages immediately while in the game (Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015). In this way, participants can rectify their actions and memorize them during the IVR SGs simulation. The second form of feedback is post-game feedback, which means participants only receive feedback after the completion of IVR SGs simulation. For instance, participants can evaluate their evacuation responses based on the final results they received after the training (Li et al., 2017). Irrespective of how it is provided, feedback is fundamental to enhancing the knowledge acquisition of participants because it will allow them to learn from their mistakes and rectify their responses to actual emergency situations.

Another teaching method was identified by this review, which is to provide instructions ahead to participants. S. Smith and Ericson (2009) applied instructions to assist participants in advance of what to do facing following up emergency situations. By doing so, it may reduce fear and stress, hence, help participants focus on training materials and complete training process since S. Smith and Ericson (2009) found that their participants (7 – 11 years old children) were nervous and fearful when facing fire emergencies without telling them what to do in their previous study. Aoki et al. (2007)

gave instructions to participants as training tours in order to help them get familiar with evacuation paths since their research was focusing on teaching spatial knowledge.

No matter which teaching method was applied, each study confirmed a positive learning outcome regarding evacuation knowledge acquisition. However, the difference in the efficiency between each method still lacks in the literature. In the reviewed papers, both teaching methods were applied only in accordance with the background of participants.

3.2.4.4.2.2. Navigation Solutions

The navigation solution is a critical factor related to a side effect of IVR SGs, commonly known as motion sickness. Participants suffer from sensory conflicts when the perceived movement inside IVR SGs does not correspond with their physical bodies' motion (Hettinger & Riccio, 1992). In this review, only one study was identified, employing a questionnaire to evaluate the motion sickness of participants (S. Smith & Ericson, 2009). It was shown that 20% of participants found motion sickness to be a disagreeable side effect during the game. However, the detailed navigation solution was not stated in the study. This study only mentioned that participants could navigate by manipulating a gamepad. Other four studies provided detailed navigation solutions (Andrée et al., 2016; Aoki et al., 2007; Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015; Duarte et al., 2014). There were two different solutions identified by this review. One was to move forward (or backward) by tilting a joystick forward (or backward), and rotate to left and right by tilting a joystick left and right (Aoki et al., 2007; Burigat & Chittaro, 2016; Duarte et al., 2014). The other was only to move towards the facing direction by holding a button on a joystick (Andrée et al., 2016; Chittaro & Buttussi, 2015). By doing this, participants need to turn their heads together with their bodies every time they wanted to turn to a new direction in the game.

The navigation solution is related to the IVR hardware. Different hardware has different representation and control system. Even for the same hardware, there are still different navigation solutions available by manipulating joysticks differently as we identified in the preceding paragraph. However, at this stage, there is no study investigating the impact of different navigation solutions on motion sickness within the scope of this study.

3.2.4.4.2.3. Narrative Methods

The narrative method refers to the method applied to encourage participants to follow and complete storylines of IVR SGs. The storyline is a set of scenarios in which participants can make decisions and take actions accordingly. In this way, the expected outcomes can be generated and achieved. The underlying reason for a specific narrative is that an IVR environment can often be simulated as an open world where participants may get lost or wander around out of curiosity (Lovreglio et al., 2018). As a result, there were four narrative methods recognized by this review, namely action-driven method (Chittaro & Buttussi, 2015), instruction-driven method (Andrée et al., 2016; Aoki et al., 2007; Burigat & Chittaro, 2016; Cosma et al., 2016; Meng & Zhang, 2014; S. Smith & Ericson, 2009; Zou et al., 2016), performance-driven method (Li et al., 2017), and surrounding-driven method (Duarte et al., 2014; Kinateder et al., 2014a; Kinateder et al., 2014b; Kinateder et al., 2015; Ronchi et al., 2015; Ronchi et al., 2016).

The action-driven method means that storylines can be driven by a sequence of actions taken by participants. Participants have to select the correct actions from very limited movement choices in order to make progress through the game (Chittaro & Buttussi, 2015). The second solution is called instruction-driven method. Prior instruction is provided to guide participants' behavior while they can still move freely in the IVR environment (Andrée et al., 2016; Aoki et al., 2007; Burigat & Chittaro, 2016; Cosma et al., 2016; Meng & Zhang, 2014; S. Smith & Ericson, 2009; Zou et al., 2016). Additionally, Burigat and Chittaro (2016) displayed instructions to bring participants back on the right track if they stopped or moved in a wrong direction. Apart from the two methods discussed above, the performance-driven method was also identified by this review. Li et al. (2017) introduced a scoring system, which can generate scores based on participants' gaming performance. By doing this, participants were encouraged to complete the storyline and do their best during the training sessions. The last method identified is called surroundings-driven method. To be specific, the IVR environment is filled with simulated hazards, leaving limited possible directions participants can move in. Eventually, the entire IVR environment will be filled with simulated hazards. Therefore, the possible movement area is so restricted that participants have no choice other than to follow and complete the pathway determined by the storyline (Duarte et al., 2014; Kinateder et al., 2014a; Kinateder et al., 2014b; Kinateder et al., 2015; Ronchi et al., 2015; Ronchi et al., 2016).

The narrative methods can largely influence on how the game is progressed by participants. Both action-driven and surroundings-driven methods make the IVR environment limited to move around, leaving participants no other choices but have to follow the storyline. In opposite, instruction-driven and performance-driven methods still keep the open world of IVR SGs. The progress of the storyline is relying on participants personal ability and willingness to complete the game.

3.2.4.4.2.4. Hazards Simulation

In the reviewed studies, several types of hazards were adopted and simulated. These hazards can be categorized into two different levels: static level and dynamic level.

The static level means that the simulated hazards are static, and they do not interact with participants; hence, they do not negatively affect participants. These hazards can be used to increase fear perception, trigger events, constrain moving area, and improve environmental realism. For instance, a burning vehicle that blocks one evacuation way (Kinateder et al., 2014b; Kinateder et al., 2015; Ronchi et al., 2015), an engine failure (Chittaro & Buttussi, 2015) or an explosion (Duarte et al., 2014) triggers the evacuation event, and high-fidelity fire simulation promotes a level of realism (Zou et al., 2016).

The dynamic level is a more advanced representation of the simulated hazards. They can interact with participants and impact participants by providing negative experience. For instance, participants can see their eyes spattered by blood if they have been hit on the head, or they can believe they are in danger of drowning if water enters the cabin of a plane (Chittaro & Buttussi, 2015). If participants are standing in an environment with smoke, they will lose visibility (Aoki et al., 2007; Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015; Cosma et al., 2016; Duarte et al., 2014; Kinateder et al., 2014a; Kinateder et al., 2014b; Kinateder et al., 2015; Ronchi et al., 2015; S. Smith & Ericson, 2009). In this case, S. Smith and Ericson (2009) suggested participants crawl on the floor to get under the smoke as the correct behavior and response. In another study (Li et al., 2017), falling and fragile objects can hurt participants' virtual bodies during an earthquake simulation. Therefore, participants need to come up with a strategy to protect themselves. No matter what type and level of hazards are chosen, participants are exposed to these realistic hazards in a completely safe environment. In this way, IVR SGs can largely influence participants' behavior.

One of the advantages of IVR SGs is that various hazards and dangerous situations, which are impossible to be presented in real life, can be presented putting no one's life at risk. On the one hand, a large number of static hazards can be applied to make progress of the game and make the IVR environment realistic. On the other hand, dynamic hazards can be adopted to create negative experience such as hurt, bleeding, or drowning in order to enhance participants' impression regarding the consequences of inappropriate responses during an emergency.

3.2.4.4.2.5. Sensory Stimulation

Another major component of the gaming environment is sensory stimulation. Four different types of stimulation were identified by this review. Obviously, every study had both visual and auditory stimulations since IVR SGs are basically video games. Apart from that, motion interaction (Li et al., 2017; S. Smith & Ericson, 2009) and olfactory stimulation (Meng & Zhang, 2014) were also recognized by this review. Li et al. (2017) introduced the motion interaction system to allow participants to physically undertake the following actions: moving, crouching, and head-protecting during the IVR SGs session. S. Smith and Ericson (2009) enabled participants to crawl physically in order to get under the smoke in the IVR SGs session. By using motion tracking systems, participants can obtain a better engagement and perception in comparison to the use of traditional joysticks or joypads. In addition, it is more user-friendly as it is easier to control and understand. Another interesting application is that Meng and Zhang (2014) proposed a smoke generator to provide olfactory and visual stimuli to participants with real smoke in compliance with the IVR SGs session. In this case, participants received a relatively high-fidelity IVR experience. The combination of different sensory stimulations is beneficial to provide a better engagement in order to minimize unrealistic behavior in the virtual environment.

3.2.4.4.2.6. Non-player Characters

Two types of NPCs were identified by this review, non-interactive NPCs and interactive NPCs. Chittaro and Buttussi (2015) introduced non-interactive NPCs only to represent other evacuees in order to improve the realism of the IVR experience. Interactive NPCs were also proposed in the same study to provide recommendations to help participants in order to achieve the intended learning outcomes. Kinateder et al. (2014a; 2014b) adopted interactive NPCs to perform social interaction with participants to investigate social influences on evacuation behavior. To be specific, these NPCs performed various actions such as running to a wrong evacuation direction so that participants' evacuation behavior can be influenced.

Representing NPCs is a challenge for the development of IVR SGs (Lovreglio et al., 2017). The reviewed papers show that IVR SGs have the possibility to investigate how participants' behavior is influenced by other evacuees. The use of NPCs can help deliver expected outcomes.

3.2.4.5. Interpreting the Findings

Based on the data analyzed above, we summarized the various factors and aspects influencing the development and implementation of an IVR SGs evacuation research into a conceptual framework, shown in Figure 3-4.

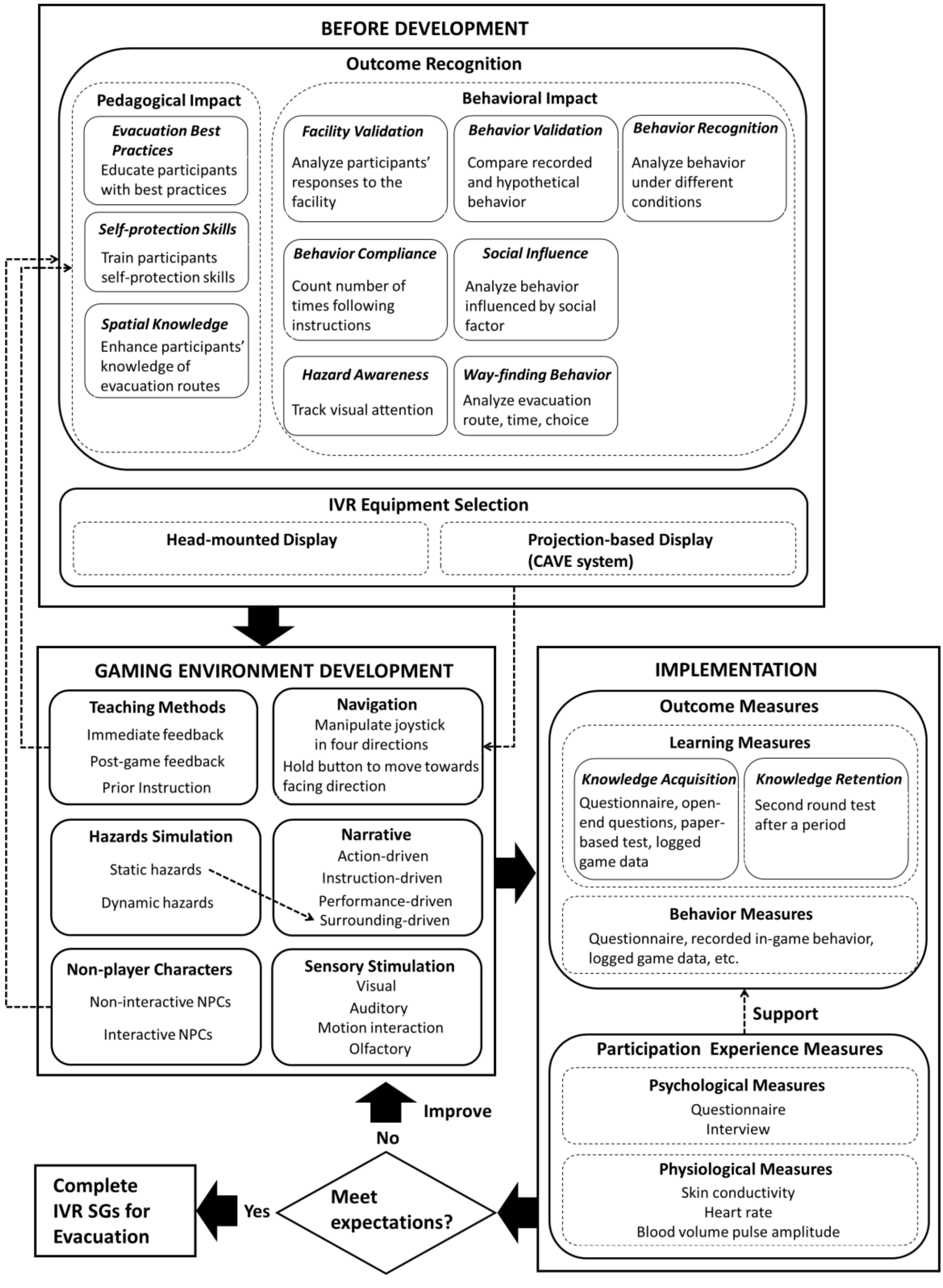


Figure 3-4 A conceptual framework for developing and implementing IVR SGs for evacuation research

Before development process, researchers need to identify expected outcomes and impacts to be achieved by IVR SGs. In order to achieve the goals of pedagogical impact, pre-defined evacuation knowledge needs to be delivered to participants effectively. There are two teaching methods that are suitable for different target groups, which are feedback after response and instruction before responses. Both methods are proven effective. The selection criteria can be based on participants' background and game narrative method. After training, multiple learning measures can be implemented to evaluate learning outcomes. The IVR SGs development also needs to include data recording function if such data (e.g., movement path, evacuation time, exit choice) is necessary for learning measures. For the behavioral impact, various outcomes can be acquired by the IVR SGs approach, such as behavior compliance, behvaior recognition, behavior validation. In accordance with different behavioral outcomes, appropriate measures need to be considered in the gaming development process in case further analysis tools are required to be added into IVR SGs in order to track behavior and record data.

Apart from the outcomes to be identified before the development, another important consideration is the IVR equipment selection. Two types of equipment were identified in this review, namely the head-mounted display (HMD) and projection-based display (PBD). Each equipment type requires different navigation solutions, which are the essential parts of IVR SGs. Appropriate navigation solutions should be designed thoroughly in accordance with the available IVR equipment in order to deliver expected outcomes by providing a comfortable gaming experience (Riecke et al., 2010).

During the development process, a few elements in the framework can be taken into consideration by developers in order to achieve high-quality IVR SGs. Simulated hazards are mainly related to simulation events. Static hazards can influence game narratives, especially the surrounding-driven method. Static hazards can make up a dangerous area and leave participants limited moving options so that they have to follow the pre-defined game storylines. Dynamic hazards can be rather flexible depending on the research purposes and participants' backgrounds. Narrative methods can significantly influence game storyline development. Participants' background, expected outcomes, and IVR environment should be taken into account when deciding the appropriate narrative method. Sensory stimulation is mainly based on available equipment. It is believed that participants should feel being physically inside the virtual world in order to minimize unrealistic behavior (Rüppel & Schatz, 2011). The use of multiple sensory stimulations can promote this influence. NPCs are also an important element of IVR SGs. Not only they can help increase the realism of IVR SGs, but also they can interact with participants to help achieve certain research goals.

The implementation phase of IVR SGs follows the development phase. During implementation, expected outcomes can be measured to validate the effectiveness of IVR SGs. Along with the outcomes measuring, participation experience can also be obtained as strong evidence to support the outcomes and to provide valuable feedback to assess the usability of IVR SGs. If the outcomes or the experiences are not satisfactory, IVR SGs need to be improved in accordance with the measured results until the results meet the expectations. After that, a final product of IVR SGs is completed, which is able to deliver the expected outcomes meeting research requirements.

3.2.5. Conclusions

We carried out a systematic literature review on IVR SGs for evacuation training and research. The pedagogical and behavioral outcomes, gaming environment development, and outcomes and participation experience measures were extensively explored. The findings indicate the advantages and disadvantages of IVR SGs in terms of delivering evacuation knowledge and conducting evacuation behavior analysis. This study provides insights into the characteristics and structure of IVR SGs. As a result, we proposed a conceptual framework for developing and implementing IVR SGs based on the investigation into the existing literature. This framework aims to contribute to future applications of IVR SGs for evacuation training and research.

When conducting this study, we found some potential directions for future research. As we stated before, there is still in need to investigate the impact of different navigation solutions on motion sickness. Apart from that, within the 15 reviewed papers, eleven focused on fire evacuation, three focused on aviation and spacecraft evacuation, while only one refers to evacuation during earthquakes. There seems to be a significant gap between the research on fire evacuation and research on earthquake evacuation, with little attention given to the latter. Regarding that, more attention needs to be paid on earthquake safety training (Lovreglio et al., 2017). Another interesting finding is that only one study took children as research subjects (S. Smith & Ericson, 2009). The rest of the researches were carried out in universities, with students and staff comprising the majority of the subjects. During an emergency, children are more vulnerable than adults. Therefore, it would be valuable to conduct more research on children using an IVR SG approach.

Chapter 4: An IVR SG for Earthquake Emergency Training

The content of this chapter is extracted from:

Feng, Z., González, V.A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., Lovreglio, R., & Cabrera-Guerrero, G. (2020). An Immersive Virtual Reality Serious Game to Enhance Earthquake Behavioral Responses and Post-earthquake Evacuation Preparedness in Buildings. *Advanced Engineering Informatics, 45*, 101118.

In Chapter 3, a general framework (see Figure 3-4) for emergency training and research has been proposed. In this chapter, a case study with Auckland City Hospital is presented, leading to the validation of the general framework.

4.1. Introduction

Earthquakes are commonly experienced disasters across the world. Every year it is estimated that 100 significant earthquakes hit different areas of the world with varying levels of structural and non-structural damage (Coburn, Spence, & Pomonis, 1992; United States Geological Survey, 2018). The structural integrity of buildings can be increased to prevent structural collapse (Ye, Qu, Lu, & Feng, 2008). Besides, proper and immediate behavioral responses during earthquakes and post-earthquake evacuation are key factors in reducing the impacts of non-structural damage (Alexander, 2012; Bernardini et al., 2016). "Drop, cover and hold" and a list of follow-on behaviors are recommended as best practice in earthquake-prone countries (Mahdavifar et al., 2009; New Zealand Ministry of Civil Defence & Emergency Management, 2015; Stuart-Black, 2015). Different educational approaches have been adopted to foster the recommended behaviors focused on building occupants such as seminars, posters, or videos. However, these educational approaches often have low emotional engagement and lack realistic hazardous situations and so may not lead to a behavioral shift towards best practice (Chittaro & Ranon, 2009). Apart from these educational approaches, building occupants also can receive practical training through evacuation drills. Bernardini et al. (2016) argued that building occupants might have different behavioral responses in evacuation drills in comparison to a real earthquake evacuation process. One possible reason is that evacuation drills are not able to realistically represent actual hazards; thus, this may lead to a reduced impact on learning outcomes and behavioral changes (Lovreglio et al., 2017). Besides, building occupants often receive no individual feedback indicating how well they conform to best practice after performing evacuation drills. Without feedback for assessment, building occupants may bring inappropriate behaviors into actual earthquake emergencies. As a result, the effectiveness of these teaching and training approaches is limited in terms of their pedagogical outcomes (Lovreglio et al., 2017).

In order to overcome the limitations mentioned above, innovative digital technologies such as Immersive Virtual Reality (IVR) and Serious Games (SGs) have been introduced for teaching and training purposes in recent years (Freina & Ott, 2015). IVR is a technology that can immerse participants in computer-generated virtual environments (LaValle, 2016). By using IVR, more realistic hazards and threats can be simulated and presented to participants in order to provide life-threatening scenarios to be used in training environments. SGs are a form of video games with a pedagogical goal as one of their primary purposes (Wouters et al., 2009). SGs can assist in the effective development of IVR educational applications. IVR SGs have been widely adopted for training, such as surgical training (Huber et al., 2017), power lines maintenance and operation training (García, Bobadilla, Figueroa, Ramírez, & Román, 2016), and pedestrian safety training (Schwebel, Combs, Rodriguez, Severson, & Sisiopiku, 2016); however, the applications for earthquake emergencies are still rare

(Feng et al., 2018).

The objective of this study is to investigate the effectiveness and applicability of IVR SGs as training tools to enhance earthquake immediate behavioral responses and post-earthquake evacuation preparedness. We propose an IVR SG framework for a hypothetical earthquake emergency occurring at Auckland City Hospital. This assisted in training participants about best evacuation practice according to the New Zealand Civil Defence guidelines (New Zealand Ministry of Civil Defence & Emergency Management, 2015) and Auckland District Health Board Evacuation Plans (Auckland District Health Board, 2009).

This paper provides the background of IVR SGs for emergency preparedness, and national and hospital earthquake response procedures in New Zealand. It then introduces the proposed IVR SG in detail, presents the research methods applied, and reports and discusses the results.

4.2. Background

4.2.1. National and hospital earthquake response procedures in New Zealand

This study took place in New Zealand, and Auckland City Hospital (ACH) was chosen as a case study. National and hospital earthquake response procedures in New Zealand are reviewed in this section.

New Zealand experiences earthquakes with a frequency of between 150 and 200 perceptible earthquakes a year (McSaveney, 2017). An M_w 6.2 earthquake hit

Christchurch in 2011, causing 185 fatalities (New Zealand Police, 2011). As a result, the New Zealand's government has put significant resources into training and educating the general public about the best evacuation practice to respond to earthquake emergencies. Since 2012, and on a yearly basis, the New Zealand's government has promoted a nationwide earthquake drill and tsunami evacuation practice named the New Zealand ShakeOut (New Zealand ShakeOut, 2018). The New Zealand ShakeOut aims to encourage the public to undertake the basic "Drop, Cover and Hold" actions during an earthquake, and to practice a tsunami evacuation if necessary. Drop, Cover and Hold (DCH) has been promoted as the primary action to perform during earthquakes in New Zealand (New Zealand Ministry of Civil Defence & Emergency Management, 2015). DCH encourages people to drop down to maintain balance, take shelter under sturdy furniture (e.g., a table) within a few steps, and hold on to it to maintain protection (Mahdavifar et al., 2009).

After shaking stops, instead of immediately exiting buildings, New Zealand Civil Defence encourages a list of behaviors representing the best practice for post-earthquake evacuation, such as check surroundings, identify suitable evacuation pathways, gather important personal items, and help others if possible (New Zealand Ministry of Civil Defence & Emergency Management, 2015). The general public has various ways to get access to learn the recommended behavioral responses to earthquakes and post-earthquake evacuation. However, they have little chance to practice it. Even in earthquake drills, building occupants often practice DCH only, not the entire set of best practice due to the cost, and the actual hazards and threats are too dangerous to be represented within a real physical setting (Becker et al., 2016; Gwynne et al., 2016; Lovreglio et al., 2017). This supports the notion that IVR SGs have potential as a training tool to promote earthquakes and post-earthquake evacuation preparedness; and therefore, they need to be further investigated.

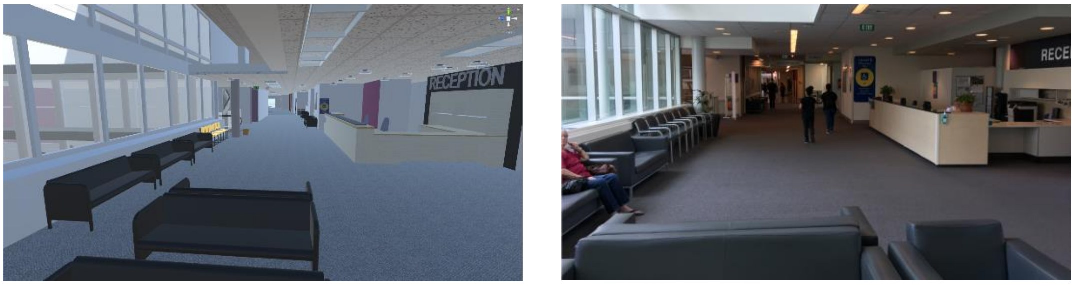
In addition to the national advice, Auckland District Health Board has issued an Auckland District Health Board Evacuation Plan (Auckland District Health Board, 2009) and an Emergency Preparedness & Response Manual (Auckland District Health Board, 2014), which include earthquake response procedures for Auckland District Health Board facilities. Similar to the recommendations suggested by New Zealand Civil Defence, this Manual encourages hospital staff to DCH during earthquakes and stay in buildings immediately after an earthquake. Also, staff members are recommended to take more responsible actions, such as administer first aid as required and advise visitors to remain until the situation has been assessed for safety. Despite having well-defined earthquake response procedures and guidelines, Wabo et al. (2012) found that there was a lack of appropriate mechanisms to effectively implement hospital evacuation plans and properly assess their impacts. In order to address this issue, Johnson (2006) suggested that hospital evacuation plans can be effectively assessed by alternative approaches such as computer simulations. Therefore, IVR SGs potentially can be investigated in order to understand how they can provide meaningful insights into the preparedness of hospital occupants during emergency evacuation; and, following that, enhance their preparedness.

4.3. The IVR SG training system

The proposed IVR SG training system allows participants to experience full indoor earthquakes and post-earthquake evacuation. The development of the IVR SG training system followed the general framework proposed in Chapter 3.

4.3.1. Virtual Environment

We chose a portion of the ACH’s fifth floor (the fourth floor is the ground floor) as the training location for our IVR SG training prototype. There were two reasons why we did not use the entire hospital building: 1. The full building model was large and complex, containing a large amount of information that could jeopardize computational performance (low frames per second) and developing processes. 2. The present study aims to teach the national evacuation guidelines in New Zealand that cover scenarios from the beginning of an earthquake to the end of post-earthquake evacuation. As long as the virtual environment was able to contain the targeted scenarios to deliver the intended training outcomes, it was not necessary to have the entire building as a training location in order to achieve the objectives sought in this research. In the case of teaching spatial knowledge, it would be necessary to include the entire building in an IVR SG framework that can be part of further research. We used a Building Information Modeling (BIM)-based workflow, which is an ideal approach to present dynamic changes such as simulating earthquakes, to develop 3D models for virtual environments (Feng, González, Ma, Al-Adhami, & Mourgues, 2018). The 3D model of this building section was developed using Autodesk Revit (a BIM tool for building modelling, www.autodesk.com). Structural components (e.g., walls, columns, floors) and non-structural components (e.g., furniture, doors, windows, ceiling tiles) were included in this BIM model. The BIM model was then imported into Unity (a game engine with user-friendly interfaces and tools for developing IVR and games, www.unity.com) for IVR and game mechanism development. More details of the workflow from BIM to IVR can be found in Lovreglio et al. (2018). Figure 4-1 compares scenes from the virtual model and the real ACH.



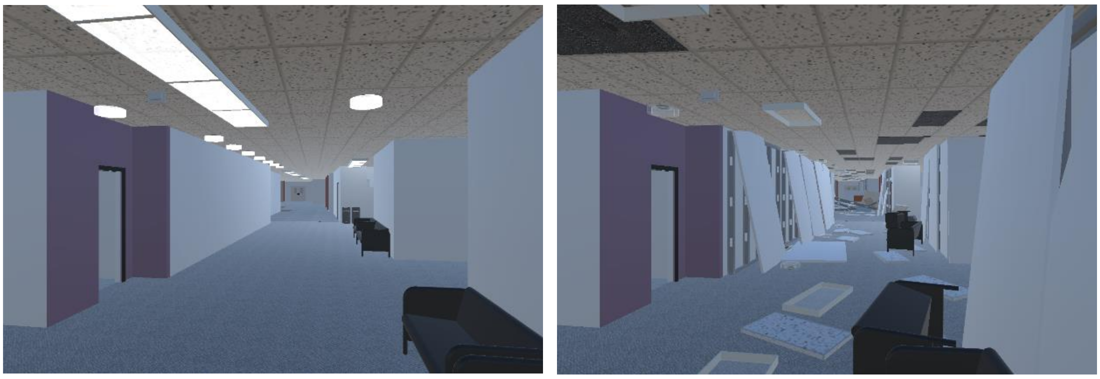
(a) The virtual reality model (b) The reality of ACH

Figure 4-1 Comparison between the virtual reality model and the reality of ACH

4.3.2. Earthquake Simulation and Building Damage

We adopted a qualitative strategy to simulate earthquake shaking and damage to the hospital building. This strategy allowed us to mimic damage based on existing datasets of videos and images of building earthquake damage, which excluded the accurate computational simulation of actual structural responses for building elements during and after earthquakes (Lovreglio et al., 2018). The reason to use this strategy is that the purpose of the IVR SG is to provide a training environment with credible and meaningful earthquake and post-earthquake experience, rather than simulations for structural analysis. In addition, the message to be sent to participants through the IVR SG is the appropriate responses to certain evacuation scenarios as recommended by the New Zealand national guidelines. Behavioral responses in a structurally damaged building are not intended by these guidelines. In this case, a major failure or collapse of structural components of the hospital building was not considered in our simulation. We only represented non-structural damage such as falling ceiling tiles, toppling partition walls and furniture, and the breaking of glass panels. Earthquake simulation and building damage were developed using Unity based on the imported BIM model. New Zealand Modified Mercalli Intensity Scale was used as the qualitative strategy data source, with the description of a severe earthquake being referred to (e.g., "Furniture and appliances are shifted. Substantial damage to fragile or unsecured objects") (GeoNet, 2019). In addition, public online videos and photos were also referred to as a complementary data source. The BIM model is a cluster of individual elements. To represent damages, we manipulated the positions and orientations of some elements, such as ceiling panels, wall panels, and furniture in Unity, providing visual cues. We also added some extra elements and models such as breaking glass in

Unity, which were not included in the original BIM model. Figure 4-2 compares a before-earthquake environment and an after-earthquake environment. Participants perceived the earthquake from visual and auditory cues.



(a) A before-earthquake environment (b) An after-earthquake environment

Figure 4-2 Comparison between a before-earthquake environment and an after-earthquake environment (Lovreglio et al., 2018)

The realism of the virtual environment and earthquake simulation and damage was measured through a self-reported questionnaire, using a 7-point Likert scale from where -3 meant strongly disagree and +3 stood for strongly agree. Participants acknowledged a high level of realism of the virtual environment (M = 1.84, SD = 1.16) along with a high level of realism of earthquake simulation and damage. Confirmatory factorial analysis indicated the presence of a latent construct, i.e., the sense of presence. The factorial analysis also provided the estimation of impacts that each factor had on the sense of presence, using a normalized scale from zero to one. Results showed that the realism of the virtual environment (0.830) and the realism of the earthquake simulation and damage (0.837) were the two major contributing factors to the sense of presence (Lovreglio et al., 2018).

4.3.3. Navigation and Interaction

We adopted a waypoint system for navigation. Waypoints are sets of coordinates that identify a stopping point or point where navigation routes can be modified (Ragavan, Ponnambalam, & Sumero, 2011). Predefined routes were used to connect the waypoints. The navigation was achieved by moving participants' view from a waypoint to another. Participants could turn their bodies to adjust the orientation of their view. This solution limited the participants' movement to prevent them from getting lost or stuck in an open-world IVR environment which occurs if they can move freely, and to reduce motion sickness from the abrupt and non-natural motion of participants (Lovreglio et al., 2018). Whenever participants reached a stopping point, they faced several options that were presented as action panels, as shown in Figure 4-3. These actions were related to the recommended behaviors as listed in Table 4-1, with some of them being appropriate actions, whereas some being inappropriate actions. Follow-on teaching methods were deployed in the IVR SG to inform participants about the recommended behaviors and facilitate learning, as discussed in Section 0. Participants could make a choice by clicking on one of the panels; and then, they proceed with the stages or scenarios of the IVR SG training system.

(a) (b)

Figure 4-3 An example of two action panels for participants to choose how to exit the building: (a) use an escalator (an inappropriate action); (b) use a staircase (an appropriate action)

4.3.4. Training Objectives and Storyline Narrative

New Zealand Civil Defence has issued a guideline providing recommendations in response to earthquakes, which includes 32 expected behavioral responses in the case of during and after an earthquake (New Zealand Ministry of Civil Defence & Emergency Management, 2015) This guideline covers behavioral responses suited to an extensive range of scenarios, such as in a building (e.g., at home, at work, or in a building away from home), outside a building (e.g., at coastal areas, at mountainous areas, or in a vehicle), or taking care of pets and livestock. We selected the behavioral responses that were feasible to implement in our designed virtual environments, focused mainly on indoor scenarios. In addition, Auckland District Health Board has an Emergency Preparedness And Response Manual (Auckland District Health Board, 2014), which covers recommendations suited to different personnel such as general staff, senior staff, and duty managers. We selected the roles that targeted general staff. The recommendations covering managerial and administrative behaviors were excluded in the present study, as they were not applicable for most of the targeted participants within the proposed structure of the IVR SG training system (occupants in a hospital include general staff and visitors). We merged the recommendations from these two sources and established a list of behaviors as training objectives shown in Table 4-1.

Table 4-1 Recommended behaviors as training objectives

Phase	Recommended Behaviors
Indoor Earthquake Phase	Drop, cover and hold
	Pay attention to falling, breaking or dangerous objects around
Pre-evacuation and Indoor Evacuation Phase	Stay under cover to see if there are aftershocks
	Collect personal belongings
	Take first aid kit

	Check and help people around
	Search for alternative exits if the closest or usual one is blocked
	Put out a small fire with a fire extinguisher or report it to the fire brigade
	Unplug damaged electrical equipment
	Use stairs to exit
Outdoor Evacuation Phase	Go to an assembly point (an open space away from buildings and falling objects)
	Do not go back to buildings until it is safe to do so

As in Feng et al. (2018), we adopted an action-driven narrative method, which means a storyline is driven by a sequence of actions taken by participants. With the waypoint system as described in the previous section, participants were led through different game scenarios in which they needed to choose actions in order to make progress through the storyline. The recommended behaviors were embedded in these game scenarios, which shaped the main storyline of the IVR SG.

The storyline of the IVR SG training system, which takes approximately 20 minutes to complete, consists of the following points, as shown in Figure 4-4:

1. Participants start the game outside ACH, and they are asked to reach a meeting room in the hospital by following a staff member.
2. Once participants have reached the meeting room, they are welcomed by a doctor Non-player Character (NPC), who invites them to leave their belongings on a table.

3. As they leave their belongings, an earthquake strikes. Participants can choose to take cover under a table, or beside a shelf or a window. If participants do nothing after ten seconds, they will be hit by a falling ceiling tile.
4. When the shaking ends, the doctor in the room leaves to check the situation outside while participants can take several actions available in the scenario, which are recommended behaviors listed in Table 4-1. Participants can take more than one actions, with no time limits. The actions included in this part are:
   i) Stay under cover to see if there are aftershocks;
   ii) Collect their personal belongings;
   iii) Take a first aid kit in the room.
5. Finally, participants have an option to get out of the room to start evacuation. While evacuating, participants come across several scenarios in which they can choose whether to take a few actions on the way out. Participants can take more than one actions, with no time limits. The actions included in this part are:
   i) Assist a nurse NPC with an injured victim;
   ii) Help a female NPC trapped under a table;
   iii) Search for an alternative exit if the closer or usual one is blocked;
   iv) Unplug damaged electrical equipment;
   v) Extinguish a small fire or report it to the fire brigade;
   vi) Listen to the radio to collect information;
   vii) Use stairs to exit;
6. Participants reach the exit of the building, and then they can choose a safe assembly point to go.
7. The experience ends in a virtual environment where participants receive a post-game assessment commenting on their behaviors. The post-game

assessment is a checklist plus a video playback with vocal commentary (see Section 0).

(a)

(b)

(c)

(d)

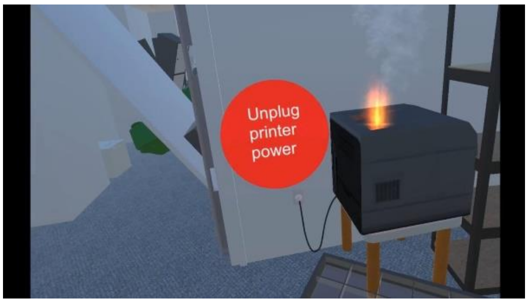
(e)

(f)

(g)

(h)

(i)



(j)

(k)

(l)

(m)

(n)

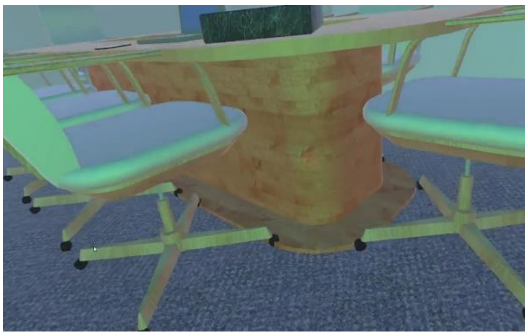
Figure 4-4 The storyline of the IVR SG training system: (a) stand outside ACH; (b) welcomed by a doctor NPC in a meeting room; (c) take cover under a table; (d) stay under cover after shaking to see if there are aftershocks; (e) collect personal belongings; (f) take the first aid kit; (g) assist a nurse NPC; (h) help a female NPC; (i) search for alternative exits; (j) unplug a damaged printer; (k) use a fire extinguisher; (l) listen to the radio; (m) use stairs; (n) go to an assembly area in an open space

The storyline was organized following the sequences of events occurring before, during, and after an earthquake. Some events might cause immediate harm to building occupants. Immediate harm was presented as a penalty effect; for instance, participants could get hit by a falling ceiling tile if they did not protect themselves when an earthquake was striking. In some other cases, such as not picking up personal belongings before evacuation, there was no associated direct harm. Participants could just ignore this action and start evacuation. In this case, no positive or negative effects were applied. However, at the end of the game, trainees still had the chance to learn what were the recommended behaviors through post-game assessment.

4.3.5. Teaching Methods

Two teaching methods were applied in this training system, namely immediate feedback and post-game assessment (Feng et al., 2018). Regarding immediate feedback, a flashing light was immediately activated after a participant made a choice, indicating whether the decision and further action were recommended or not. If participants chose a recommended action, green lights flashed; whereas for an action that was not recommended, red lights flashed. In this way, participants could immediately receive feedback related to the assessment of their actions. Figure 4-5 shows an example of flashing green lights indicating participants have chosen a recommended action, and flashing red lights indicating an action that is not recommended.

(a)



(b)

(c) (d)

Figure 4-5 An example of flashing lights: (a) crouch under a table when an earthquake hits, (b) green lights flash; (c) crouch beside a window, (d) red lights flash

Once the entire evacuation process was over, participants received a detailed post-game assessment, which reported all the actions that had been taken against the full list of recommended behaviors listed in Table 4-1. Following that, a video and audio playback took participants through all the choices that they made during the training experience, explaining the rationale behind each recommended behavior, as shown in Figure 4-6. The post-game assessment served as a recap to help participants understand what the recommended behavior was and strengthened their memory to reinforce them.

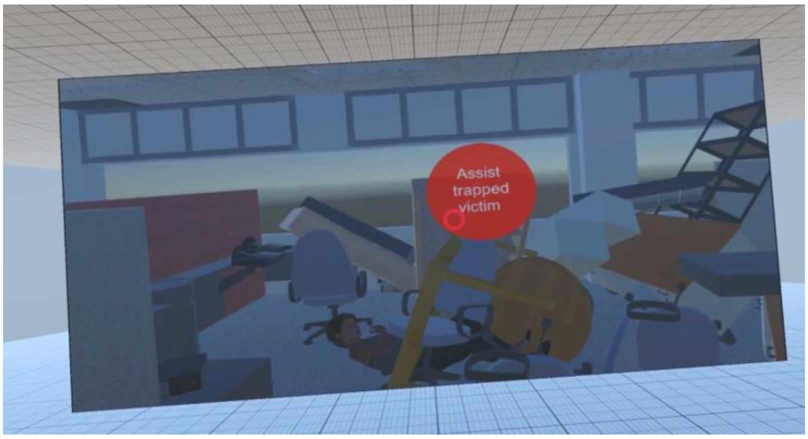


Figure 4-6 A screenshot of playback explaining the recommended behaviors

4.4. Research Methods

To evaluate the possible effectiveness and applicability of the IVR SG training system, we chose ACH as a case study. ACH is New Zealand's largest public hospital and clinical research facility. The reason to use ACH in the IVR SG training system is that hospital evacuation drills are always restricted due to ethical issues and risks from disruptions of operational functions (C. W. Johnson, 2006). This gave us the opportunity to run a virtual drill by using IVR SGs. We used a one-group pretest-posttest research design. We carried out a pre-test measure of the outcome of interest prior to administering training, followed by a post-test on the same measure after training occurred. Participants received training through the IVR SG training system, mainly focusing on the behavioral responses to indoor earthquakes and post-earthquake evacuation. This study did not include a control group with a traditional training method, as the aim of this study was to evaluate the effectiveness and applicability of IVR SGs for earthquake emergency training. A further discussion of the limitation of this study design is made in Section 8.3.

4.4.1. Apparatus

The IVR SG training system was implemented as an executable file which was built in Unity. The IVR SG training system was run on a DELL PC workstation equipped with a 2.4 GHz Intel Xeon E5-2640 processor, 64 GB RAM, and two NVidia GTX 1080 graphic cards. The IVR headset was an Oculus Rift, which is a head-mounted display (HMD) with 1080x1200 resolution per eye and a 110-degree field of view (both horizontal and vertical) (Oculus, 2020). The remote controller was an Oculus Remote, which was given to participants to choose action panels by simply pressing one button. The graphic output of the HMD was also displayed on an LED screen, which allowed researchers to observe participants' in-game behaviors during the training. The record of these behaviors (i.e., their in-game choices) was done within the IVR SG training system.

4.4.2. Participants

A total of 93 participants (43 males, 50 females) were recruited to test the IVR SG training system. Participants were recruited through emails, notices posted on the Auckland District Board staff newsletters, and leaflets and posters spread through ACH and The University of Auckland. The staff members of ACH and the general public recognized as the visitors of ACH were included as our participants. There were no exclusion criteria applied to participants. Of these, 87 participants completed the experiment. The other six had to stop the IVR experience due to motion sickness. The remaining 87 participants consisted of 25 staff members of ACH and 62 visitors. Ages ranged from less than 20 to over 70, with one-third between 20 to 29. The demographic profile is shown in Figure 4-7.

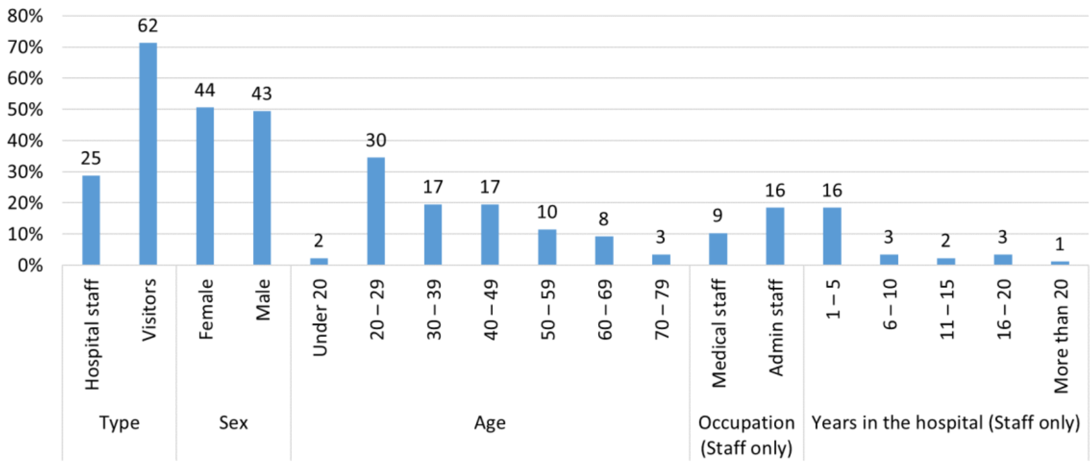


Figure 4-7 The demographics of participants

Apart from demographic information, we asked participants about their previous experience involving fire and earthquake drills. Table 4-2 shows the result of this survey revealing that participants had more experience with fire drills than earthquake drills. The possible reason is that in New Zealand, regulations require that building owners maintain an evacuation scheme by conducting fire drills or training and assessing the capability of the permanent occupants to manage evacuation at least every six months (The New Zealand Parliament, 2018), whereas there are no requirements for mandatory earthquake drills. Our participants might not be familiar with the procedure of earthquake evacuation.

Table 4-2 Frequencies of practice in fire drills and earthquake drills

Frequency	Staff		Visitors	
	Fire	Earthquake	Fire	Earthquake
Never	7	22	10	45
Once a year	8	1	26	8
Twice a year	4	0	11	0
More than twice a year	3	1	9	1
Unsure	3	1	6	5
Other	0	0	0	3 (Only once)

We also asked participants to state how often they play video games, if at all. Table 4-3 shows the result of this survey. Most of the participants were not used to playing video games often.

Table 4-3 Frequencies of playing video games

Frequency	Staff	Visitors
Never	11	19
Less than once a year	3	17
At least once a year	5	9

At least once a month	2	4
At least once a week	1	5
Several days a week	2	3
Every day	1	5

Finally, we asked participants to state whether they had experienced IVR (e.g., games, videos, tours, or demos) before. Table 4-4 shows the result of this survey. Over half of the participants had never experienced IVR before.

Table 4-4 Experience with IVR

Experience	Staff	Visitors
No	13	31
Yes	11	23
Unsure	1	8

The adequacy of the sample size was assessed. According to meta-analysis studies on Virtual Reality Exposure Therapy (Carl et al., 2019; Morina, Ijntema, Meyerbröker, & Emmelkamp, 2015), the effect size of IVR training for the entire population was medium to large. In this study, Cohen's d (i.e., the effect size) was estimated to be 0.5 (J. Cohen, 2013). A power analysis using the statistical software G*Power 3.1 indicated that a total sample of 35 people would be needed to detect a medium effect (d = 0.5) with 80% power using a t-test. A total sample of 56 people would be needed to detect a medium effect (d = 0.5) with 95% power using a t-test. In addition, Gall et al. (2006) suggest a sample size of larger than 15 cases for causal-comparative and experimental methodologies. The sample sizes of other IVR studies with similar methodologies fall in the range from 16 to 59 for each sample group (Li et al., 2017; Lin, Cao, & Li, 2019;

Lin et al., 2020). In this study, it is argued that the sample size is adequate for the experiments undertaken; however, the author is still aware that large scale experiments with random sampling are necessary to generalize the findings to a larger population (L. Cohen, Manion, & Morrison, 2013).

4.4.3. Measures

We use the preparedness of individuals as an indicator of effectiveness assessment. Individuals' preparedness for earthquakes and post-earthquake evacuation can be influenced by various factors such as hazard perception, attitude, and experience (Tekeli-Yeşil, Dedeoğlu, Tanner, Braun-Fahrlaender, & Obrist, 2010). This study deployed two instruments to assess participants' preparedness: 1) Knowledge about best evacuation practice; and 2) Self-efficacy in dealing with earthquake emergencies. A knowledge test and a self-efficacy questionnaire were both answered before and immediately after the training in order to measure the effect on the preparedness of individuals. In addition to effectiveness, we assess the applicability of the IVR SG training system by conducting a questionnaire after the training. Applicability refers to the capability of a product applied to target groups (Rhebergen, Hulshof, Lenderink, & van Dijk, 2010). In this study, applicability has been assessed using three dimensions: 1) Ease of use; 2) Training efficacy; 3) Engagement (Chittaro & Sioni, 2015). Ease of use means that the training system was easy to understand and use; training efficacy refers to the usefulness of the training content in enhancing individuals' preparedness; engagement measures the extent to which the IVR SG is able to attract the attention of participants (Chittaro & Sioni, 2015).

4.4.3.1. Knowledge acquisition

In order to measure participants' knowledge associated with the immediate behavioral responses to earthquakes and post-earthquake evacuation, we asked five questions. These five questions were focused on three aspects, namely knowledge of behavioral responses: inside a building during an earthquake, inside a building after an earthquake, and outside a building after an earthquake. The questions were open-ended questions in order to avoid the suggestion of possible answers that could bias responses from participants. Participants answered the same five questions orally before and immediately after the training, and their answers were audio recorded. Each recorded audio was transcribed and coded by three researchers. For each assessed aspect, three transcriptions from three researchers were cross-checked and merged into one final transcription. According to the final transcriptions, scores were given based on a knowledge scale. The knowledge scale was developed based on the recommended behaviors that were identified as training objectives in Table 4-1. Possible scores ranged from 1 to 4, where 1 stood for no knowledge and 4 stood for strong knowledge. As a result, every participant received three scores for pre-training and three scores for post-training. Table 4-5 shows the assessed aspects, open-ended questions, and knowledge scale.

Table 4-5 Assessed knowledge aspects, open-ended questions for pre- and post-training, and knowledge scale

Knowledge aspects	Asked questions	Knowledge scale			
		Strong knowledge (4 points)	Adequate knowledge (3 or 3.5 points)	Weak knowledge (2 or 2.5 points)	No knowledge (1 point)
Behavioral responses inside a building during an earthquake	Q1: What would you do during an earthquake? Q2: What should you pay attention to during an earthquake?	4 points for knowing to (i) drop, cover and hold under a table, (ii) pay attention to falling or unsteady objects and glass.	3 points for knowing to drop, cover and hold under a table.	2 points for knowing to take cover or find a shelter.	1 point for knowing nothing.
Behavioral responses inside a building after an earthquake	Q3: What would you do after an earthquake? Q4: What should you pay attention to after an earthquake?	4 points for knowing over eight items out of eleven items as stated below (i) stay undercover to see if there are aftershocks, (ii) collect personal items, (iii) pay attention to first aid kit, (iv) pay attention to people around and offer help, (v) search for available exits if common ones are blocked, (vi) pay attention to fire, (vii) pay attention to fire extinguishers, put out fire if practicable or call fire departments, (viii) pay attention to electric leakage, (ix) unplug equipment if practicable, (x) listen to a radio to get more information and instructions, (xi) use stairs to exit buildings.	3 points for knowing five or six items as specified in the strong knowledge column; 3.5 points for knowing seven or eight items.	2 points for knowing one or two items as specified in the strong knowledge column; 2.5 points for knowing three or four items.	1 point for knowing nothing.
Behavioral responses outside a building after an earthquake	Q5: What is the correct behavior when you are outside a building after an earthquake?	4 points for knowing to (i) stay at an open space which is away from buildings and falling objects, (ii) don't go back to buildings until it's safe to do so.	3 points for knowing to stay at an open space which is away from the buildings and falling objects.	2 points for knowing to go to an assembly point only.	1 point for knowing nothing.

4.4.3.2. Self-efficacy

Self-efficacy is a person's belief in his or her ability to successfully accomplish difficult tasks (Bandura, 1982). Self-efficacy can largely influence a person's behavior and performance outcomes (Bandura, 1977; Stajkovic & Luthans, 1998). In order to measure participants' levels of self-efficacy in dealing with earthquake emergencies, we administered a questionnaire before and immediately after the training. The questionnaire was designed based on the General Self-Efficacy Scale (Schwarzer & Jerusalem, 2010). We asked participants to rate their levels of agreement on a 7-point Likert scale (-3 = strongly disagree, +3 = strongly agree) about six statements:

1. "I am confident that I am able to effectively deal with an earthquake emergency";
2. "Thanks to my resources, I know how to manage in an earthquake emergency";
3. "I would be able to deal with an earthquake emergency even if the building is severely damaged";
4. "I would be able to deal with an earthquake emergency even if I find flame and fire along the way";
5. "I would be able to deal with an earthquake emergency even if the exit is blocked";
6. "I would be able to deal with an earthquake emergency even if I find objects that may harm me along the way."

4.4.3.3. Ease of use

In order to facilitate applicability, one important point is to ensure the training is easy to follow. We administered a set of questions immediately after the training to measure the levels of ease of use perceived by participants. The wording of the questions was based on the questionnaire used by Chittaro and Sioni (2015). We asked participants to rate their levels of agreement on a 7-point Likert scale (-3 = strongly disagree, +3 = strongly agree) with regard to three statements:

1. "I could easily learn the recommendations provided in the virtual game";
2. "I could easily remember the recommendations provided in the virtual game";
3. "I could easily carry out the recommendations provided in the virtual game."

4.4.3.4. Training efficacy

Training efficacy refers to the usefulness of the training content in reducing risks related to hazards (Chittaro & Sioni, 2015). We administered a set of questions immediately after the training to measure the levels of training efficacy perceived by participants. The wording of the questions was based on the questionnaire used by Chittaro and Sioni (2015). We asked participants to rate their levels of agreement on a 7-point Likert scale (-3 = strongly disagree, +3 = strongly agree) with regard to three statements:

1. "The recommendations provided in the virtual game are useful for my safety";
2. "The provided recommendations will allow me to effectively deal with an earthquake emergency";
3. "By following the recommendations provided in the virtual game, I can strongly reduce the probability of injury to myself or others during an earthquake emergency."

4.4.3.5. Engagement

To assess to what extent the training system attracted participant's attention, we administered a questionnaire immediately after the training. We asked participants to rate their levels of agreement on a 7-point Likert scale (-3 = strongly disagree, +3 = strongly agree) about this statement: "The game was engaging/fun."

4.4.4. Procedure

The experiment was carried out in a meeting room at ACH. Ethical clearance was granted from The University of Auckland Human Participants Ethics Committee, and the reference number was 016763. The procedure of the experiment is illustrated in Figure 4-8. Participants received participation information sheets, which informed them that this experiment aimed to test an IVR SG training system that was designed for an earthquake emergency. Participants then received consent forms, which requested consent for participation and collecting research data, including questionnaires, audio recordings, and in-game actions recordings by the researchers. Participants were informed that they could stop and quit the experiment at any time without giving any reason.

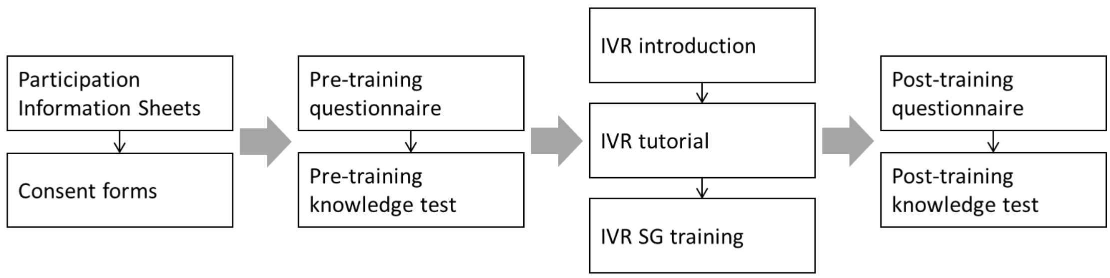


Figure 4-8 The experimental workflow

After signing the consent forms, participants were asked to fill in a questionnaire including demographic information, frequency of practice in fire drills and earthquake drills, frequency of playing video games, experience with IVR, and self-efficacy in dealing with an earthquake emergency. Following that, participants were asked to orally answer a five-question knowledge test for pre-training knowledge assessment.

After pre-training knowledge assessment, participants were given introductions about controls as well as health and safety for using IVR. Then, participants were invited to wear an HMD and adjust it until they had a clear view and felt comfortable with it. Participants were asked to sit in a swivel chair for the entire IVR SG training session, which made it easy to turn their bodies with minimum falling risk. Before participants were exposed to the IVR SG training session, they were led through a tutorial session (an empty room with waypoints and action panels) that helped participants understand the navigation and interaction of the IVR SG training system and get familiar with IVR environments and controls. Then, the training session was started. The entire training session was around 20 minutes.

After the IVR SG training session was completed, participants orally answered the same five-question knowledge test again for post-training knowledge assessment. Then, participants filled in questionnaires about self-efficacy, training efficacy, and engagement.

Before the experiment, participants were informed by Participation Information Sheet that the experiment involved visual simulations using IVR, in which case normal vision was essential; however, it was still possible to use personal glasses. There were cases that participants could fit their glasses in IVR headsets. For those whose glasses were too big for IVR headsets, participants were assisted in having a trial without glasses. If they could not get a clear view in the IVR environment, they had to drop the experiment.

4.5. Results

Two groups of participants were identified by this study, namely staff and visitors. Staff members were expected to be more familiar with the building and have a greater preparedness for emergencies. Therefore, the results of these two groups were analyzed separately as differences might be recognized.

4.5.1. Knowledge acquisition

To assess whether participants acquired knowledge using the IVR SG training system, we compared their knowledge scores before and after the training for both staff and visitors. The boxplots in Figure 4-9 show that the average knowledge on best evacuation practice has increased for staff and visitors, considering the three behavioral response categories. To investigate if these increments are significant, we used the statistical tests described below.

Given that these scores are non-normally distributed based on Shapiro-Wilk tests, Wilcoxon Signed-rank tests were used to assess whether both staff and visitors increased their knowledge by using the IVR SG. The tests indicated that both staff and visitors significantly increased their knowledge after the training, as shown in Table 4-6.

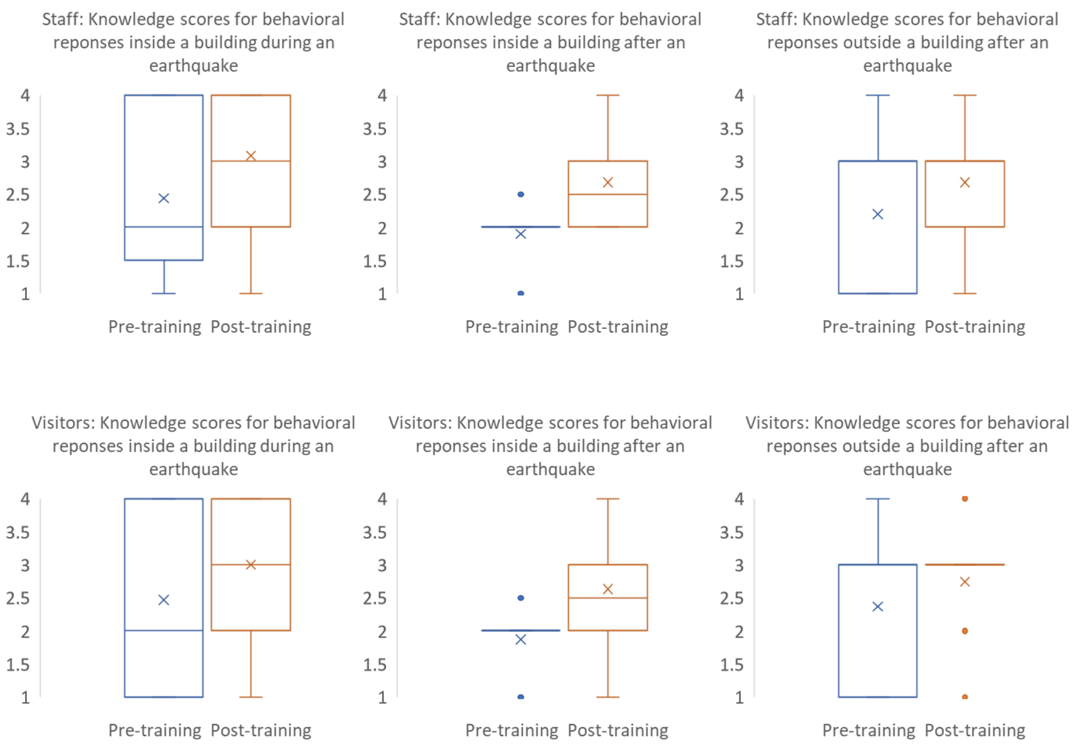


Figure 4-9 Comparisons of participants' knowledge scores using boxplots

Table 4-6 Wilcoxon Signed-rank test results for knowledge levels comparisons

Knowledge Aspects	Staff		Visitors	
	Pre-training	Post-training	Pre-training	Post-training
Behavioral responses inside a building during an earthquake	M = 2.44 SD = 1.16 p = 0.014	M = 3.08 SD = 1.00	M = 2.47 SD = 1.26 p = 0.002	M = 3.00 SD = 0.99
Behavioral responses inside a building after an earthquake	M = 1.90 SD = 0.43 p < 0.001	M = 2.68 SD = 0.64	M = 1.87 SD = 0.46 p < 0.001	M = 2.63 SD = 0.57
Behavioral responses inside a building after an earthquake	M = 2.20 SD = 1.08 p = 0.030	M = 2.68 SD = 0.63	M = 2.37 SD = 0.94 p = 0.003	M = 2.74 SD = 0.68

4.5.2. Self-efficacy

To assess the self-efficacy of participants, we asked them to provide a level of agreement to six statements. The self-efficacy scores for each participant before and after the use of the IVR SG were calculated using factorial analysis (the maximum likelihood method was used for extraction, and a regression method was used to compute factor scores). A Cronbach's alpha was also calculated to assess the internal consistency of these six statements (staff pre-training: 0.91, post-training: 0.96; visitors pre-training: 0.95, post-training: 0.93). Figure 4-10 shows the scores obtained by staff and visitors before and after the training, and there is an increment for both groups.

Given that these scores are non-normally distributed based on Shapiro-Wilk tests, Wilcoxon Signed-rank tests were used to compare the pre- and post- scores. The tests indicated that both staff and visitors had a significant increment of self-efficacy after the training, as shown in Table 4-7.

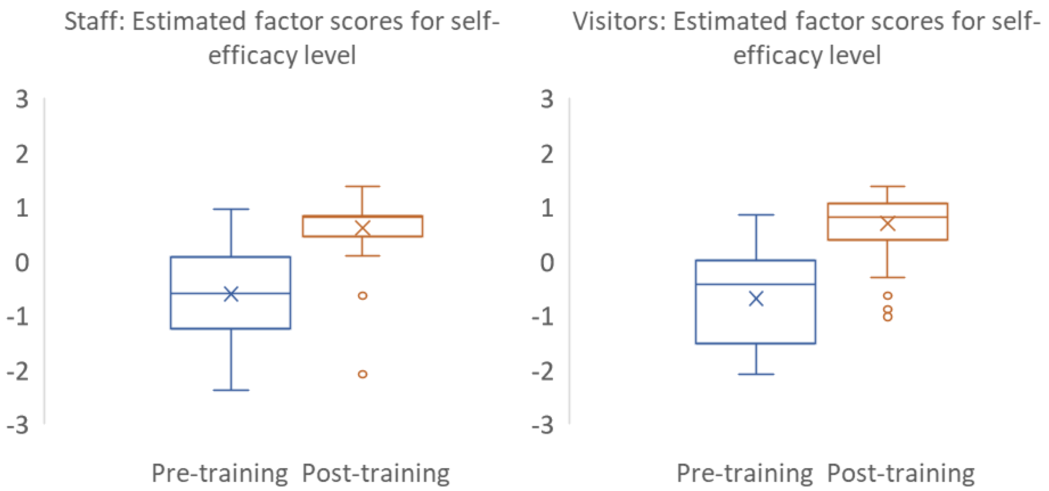


Figure 4-10 Comparisons of self-efficacy levels using boxplots

Table 4-7 Wilcoxon Signed-rank test results for self-efficacy levels comparisons

Staff		Visitors	
Pre-training	Post-training	Pre-training	Post-training
M = -0.61	M = 0.61	M = -0.69	M = 0.69
SD = 0.85	SD = 0.70	SD = 0.94	SD = 0.59
p < 0.001		p < 0.001	

4.5.3. Ease of use

Three scores were provided by each participant to assess the ease of use. The final score for each participant was calculated by averaging these scores. A Cronbach's alpha was also calculated to assess the internal consistency of these three statements (Staff: 0.87, Visitors: 0.88). The result of ease of use was reported as boxplots (Staff: M = 2.57, SD = 0.56; Visitors: M = 2.52, SD = 0.67), as shown in Figure 4-11. The result shows that participants felt that it was easy for them to learn from the IVR SG.

Figure 4-11 Perceived ease of use by participants using boxplots

4.5.4. Training efficacy

Three scores were provided by each participant to assess the training efficacy. As we did before, the final score for each participant was calculated by averaging these scores. A Cronbach's alpha was also calculated to assess the internal consistency of these three statements (Staff: 0.67, Visitors: 0.83). The result of training efficacy was reported as boxplots (Staff: M = 2.35, SD = 0.72; Visitors: M = 2.40, SD = 0.69), as shown in Figure 4-12. The result shows that participants felt that the IVR SG was useful and helpful.

Figure 4-12 Perceived training efficacy by participants using boxplots

4.5.5. Engagement

One score was provided by each participant to assess the engagement. The result of engagement rating was reported as boxplots (Staff: M = 2.28, SD = 0.79; Visitors: M = 1.79, SD = 1.26), as shown in Figure 4-13. The result shows that participants perceived that the IVR SG was engaging.

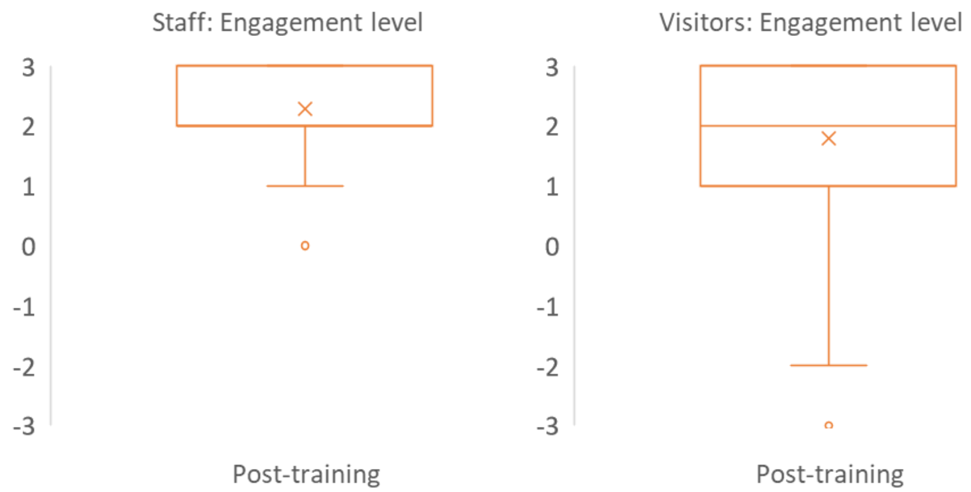


Figure 4-13 Perceived engagement level by participants

4.6. Discussion

This study provides insights into the effectiveness and applicability of IVR SGs used to improve individuals' preparedness for indoor earthquakes and post-earthquake evacuation. The results of the knowledge test in Figure 4-9 show that participants had a significant increase in knowledge about behavioral responses during earthquakes and post-earthquake evacuation immediately after the training. The results of self-efficacy measurement in Figure 4-10 indicate that participants had a significant improvement in their confidence about their ability to deal with earthquakes and post-earthquake evacuation after the training. Taken together, these results suggest that the proposed IVR SG training system is effective in increasing an individual's preparedness for earthquake emergencies. The results of ease of use, training efficacy and engagement measurements in Figure 4-11, Figure 4-12, and Figure 4-13 reveal that the IVR SG training system is applicable and engaging, and it facilitates teaching behaviors.

Emergency training targeting various emergency types have different levels of training workloads for participants. Studies using IVR SGs have allowed for the training of single behavioral responses such as self-protection skills during earthquakes (Li et al., 2017) and multiple behavioral responses such as the best practice for aviation emergencies (Chittaro & Buttussi, 2015). Our study introduces 13 behavioral responses based on the recommendations from New Zealand Civil Defence and Auckland District Health Board, which were identified as training objectives for the IVR SG training system, as shown in Table 4-1. These behavioral responses made up earthquakes and post-earthquake evacuation procedures in three phases, namely: indoor earthquake phase, pre-evacuation and indoor evacuation phase, and outdoor evacuation phase. Participants increased their knowledge significantly after the training in all three phases. This finding is consistent with that of Chittaro and Buttussi (2015) who pointed out that their IVR SG was effective in increasing safety knowledge, including ten behavioral responses on aviation emergencies. Taken together, these findings show that IVR SGs have the potential to effectively deliver complex training and learning content in terms of the best practice for emergency responses.

While IVR SGs are suggested to be effective to train participants about multiple behavioral responses with certain workloads, different performances in our training outcomes are apparent. As shown in Table 4-8, knowledge of behavioral responses inside a building during an earthquake achieved the highest scores both in pre- and post-training as compared to the other two assessed knowledge aspects. One possible factor that can contribute to this is that we assessed eleven items in knowledge on behavioral responses inside a building after an earthquake. Participants might give different degrees of detail when answering knowledge test questions since open-ended questions heavily rely on the effort from respondents; therefore, this could lead to an inaccurate assessment (Vinten, 1995). Another possible explanation for this could be that participants had a lack of awareness of what to do immediately after earthquakes. Which is evidenced by other recent earthquakes, where behaviors like freezing-in-place (as opposed to drop, cover, hold) are still the most common behavior (Lindell et al., 2016). According to the survey on the survivors of real-world earthquakes conducted by Lindell (2016), people have no idea about what to do immediately after an earthquake. This freezing-in-place behavior could be triggered by immersion and the high level of realism of earthquake dynamics and building damage, including visual and auditory simulation. As reported in Section 4.3.2, participants acknowledged the realism of earthquake simulation and damage. In our study, participants need more effort to gain and retain knowledge about the behavioral responses immediately after an earthquake. In order to address this issue, we believe that multiple practices with different training environments can be applied to participants. As Steven (1982) argued, memory recall can be enhanced by using multiple environmental contexts during learning processes. This is aligned with Chittaro and Sioni's (2015) suggestions that repetitive rehearsals can be introduced to improve learning outcomes of IVR SGs.

Table 4-8 Pre- and post-training knowledge scores for all participants

Knowledge Aspects	Pre-training	Post-training
Behavioral responses inside a building during an earthquake	M = 2.46 SD = 1.23	M = 3.02 SD = 0.99
Behavioral responses inside a building after an earthquake	M = 1.88 SD = 0.45	M = 2.64 SD = 0.59
Behavioral responses outside a building after an earthquake	M = 2.32 SD = 0.98	M = 2.72 SD = 0.66

In Section 4.5.1, we split knowledge acquisition analysis between staff and visitors. According to ANCOVA tests, there were no significant differences identified between these two groups (see Table 4-9). One possible reason is that the national guidelines also cover the selected behaviors identified as the training objectives of the IVR SG training system from the ACH evacuation plan. We did not explicitly include particular behaviors in our IVR SG training system for hospital staff. The training content was generally applicable to both staff and visitors. In addition, the familiarity of the building (where staff are likely to be more familiar with the building) might not be a contributor for knowledge gain per se, as participants were expected to acquire safety knowledge rather than perform evacuation based on their spatial knowledge of the building. In the case where evacuating as soon as possible is the priority (e.g., in fire emergencies), spatial knowledge is essential. Another possible reason is the difference in sample size (25 vs. 62), which might weaken possible differences between staff and visitors.

Table 4-9 ANCOVA tests for between-group difference analysis

Knowledge aspects	ANCOVA results
Behavioral responses inside a building during an earthquake	$F(1, 84) = 0.164$, $p = 0.686$
Behavioral responses inside a building after an earthquake	$F(1, 84) = 0.080$, $p = 0.778$
Behavioral responses outside a building after an earthquake	$F(1, 84) = 0.021$, $p = 0.885$

Apart from knowledge acquisition, our study also measured self-efficacy as another factor to influence earthquakes and post-earthquake evacuation preparedness. Self-efficacy is a personal belief and an important predictor of attitude and behavior change (Bandura, 1977). Our study shows that participants increased their self-efficacy significantly after the training. This finding is in agreement with Chittaro and Sioni's (2015) findings, which showed that SGs were effective in increasing individuals' self-efficacy in terror attack emergency preparedness. Chittaro and Sioni (2015) argued that SGs provided effective actions for participants to choose when they were threatened by risks. In this way, it was beneficial to increase self-efficacy since participants were motivated to take actions to protect themselves (Chittaro & Sioni, 2015). IVR SGs have the potential to provide engaging environments for participants to go through life-like hazards and be trained to respond to them effectively. As a result, participants are motivated to be more confident facing actual earthquake emergencies and to perform better when dealing with such emergency situations (Stajkovic & Luthans, 1998).

In terms of applicability, participants reported a high level of perceived ease of use, training efficacy, and engagement. This finding indicates that participants felt engaged, and the knowledge learning process was easy and helpful with the IVR SG training system, which is in accordance with previous studies (Chittaro & Sioni, 2015; S. Smith & Ericson, 2009). Participants from Chittaro and Sioni's (2015) were students from universities (Mean age = 23.68), while participants from S. Smith and Ericson's (2009) were children aged from seven to eleven. In our study, 87 final participants were mainly adults, with eleven of them being between 60 to 79 years old. Before the study, our participants were not familiar with gameplay and IVR, in accordance with the responses collected regarding their prior experiences with video games and IVR. Surprisingly, as an innovative digital technology, IVR SG seems to be well accepted and easy to use as an emergency training tool across various age groups This is especially important for children at school age as they are receptive to new knowledge and play an important role for a community to build up disaster prevention culture, as highlighted in previous studies (Bernhardsdottir, Musacchio, Ferreira, & Falsaperla, 2016; Izadkhah & Hosseini, 2005; Shaw, Kobayashi, & Kobayashi, 2004).

4.7. Conclusions

There has been little research on IVR SGs for earthquakes, especially focused on the full behavioral responses for earthquakes and post-earthquake evacuation. In order to fill this gap, we investigated the effectiveness and applicability of an IVR SG as a training tool to enhance the immediate behavioral responses to earthquakes and post-earthquake evacuation preparedness. We have shown that the proposed IVR SG training system was effective to enhance an individual's preparedness for earthquakes and post-earthquake evacuation. Participants' knowledge of behavioral responses and self-efficacy increased significantly after the training. Additionally, we have shown that the IVR SG was engaging and easy to use for learning immediate behavioral responses to earthquakes and post-earthquake evacuation. This demonstrates that although novel, IVR SG has the potential to be applied as a robust tool for emergency response training.

Potential directions for future studies are also identified. Firstly, as we discussed before, future studies can pay attention to the comparison of IVR SGs with traditional training methods regarding earthquake safety. Secondly, knowledge retention is also worth assessing. Thirdly, challenges may exist when participants recall their knowledge in a different environment (or context) and transfer their knowledge to actual actions. Li et al. (2017) tested their participants in a different IVR environment after learning self-protection skills in an IVR SG. Results indicated that participants could recall what they learned and apply it to the new environment. In line with the concept of situated learning (Stein, 1998), future research can extend the current IVR SG with the incorporation of repeated exercises in expansive contexts, which could contribute to the transferring of knowledge and applying of skills to new settings (Engle, Lam, Meyer, & Nix, 2012; National Research Council, 2000). In addition, future research can explore the effect of the manipulation of narratives (i.e., narrative methods), as storylines and narratives help participants with the understanding and recall of presented content (Bransford, Sherwood, Hasselbring, Kinzer, & Williams, 2012). Fourthly, we deployed green and red lights as immediate feedback. In the case of extending the current research to a large population, we can explore other forms and mechanisms of feedback to avoid the issues caused by color blindness such as text or signs. Lastly, most participants in our study were young adults. How IVR SGs fit for children and seniors in terms of earthquakes and post-earthquake evacuation training remains unclear.

This chapter validates the general framework (see Figure 3-4) proposed in Chapter 3. The pedagogical impacts and applicability of an IVR SG have been evaluated in this chapter. In the next chapter, a customization framework is built based on the general framework.

Chapter 5: A Customization Framework for IVR SGs Suited to Earthquake Emergency Training

The content of this chapter is extracted from:

Feng, Z., González, V.A., Mutch, C., Amor, R., Rahouti, A., Baghouz, A., Li, N., & Cabrera-Guerrero, G. (2020). Towards a Customizable Immersive Virtual Reality Serious Game for Earthquake Emergency Training. *Advanced Engineering Informatics, 46*, 101134

A general framework (see Figure 3-4) for the development and implementation of IVR SGs has been proposed and validated in Chapter 3 and 4. In this chapter, a customization framework for IVR SGs suited to earthquake emergency training is developed based on the integration of the general framework (see Figure 3-4) and the concept of adaptive game-based learning. In addition, a case study with children and adults is presented to evaluate the usability of a customizable IVR SG.

5.1. Introduction

Earthquakes are catastrophic events that take many human lives every year. Appropriate behavioral responses during and after earthquakes play an important role in reducing casualties and injury (Alexander, 2012; Bernardini et al., 2016). Different educational and training approaches have been applied to equip people with the most appropriate safety knowledge of best evacuation practices. As well as traditional approaches, such as drills and seminars, Immersive Virtual Reality Serious Games (IVR SGs) have increasingly gained popularity for training people (Feng et al., 2018). IVR SGs engage trainees through three core mechanisms: immersion, interactivity, and user involvement (Freina & Ott, 2015). IVR SGs offer life-like scenarios involving hazards, where trainees interact with in-game objects and solve challenges. Chittaro and

Buttussi (2015) argue that IVR SGs enhance the perception of trainees through high-level physiological and emotional arousal. Krokos et al. (2019) claim that IVR SGs help trainees focus on tasks that lead to improved memory recall. However, the existing training programs that use IVR SGs are mostly concentrated on ad-hoc scenarios or single trainee types. These IVR SGs have limited flexibility to be adapted to heterogeneous trainee types. The flexibility of SGs means it is possible to alter the game elements such as content, user interfaces, game difficulty, and game mechanisms (Streicher & Smeddinck, 2016). A fixed training program that does not take different trainee backgrounds and goals into account could jeopardize the effectiveness of the training outcomes, as trainees have different learning styles, objectives, and capabilities (Kelley, 1969).

In recent years, in order to address the limitations described above, adaptive game-based learning has been introduced (Hwang, Sung, Hung, Huang, & Tsai, 2012). Adaptive game-based learning engages trainees on a one-to-one basis. Adaptation is a process that recognizes trainee needs and backgrounds and shapes the training to be a better fit for trainees (Lopes & Bidarra, 2011b). Adaptation in SGs can occur in various dimensions, including learning content, narratives, scenarios, game worlds, and pedagogical interventions (Kickmeier-Rust & Albert, 2010; Lopes & Bidarra, 2011a). With the integration of the adaptation idea, IVR SGs can more effectively deliver training outcomes by generating a personalized training experience. Such a personalized experience is beneficial to inducing high engagement and in-depth cognitive information-processing (Teng, 2010). To date, little research has paid attention to adaptive IVR SGs, especially for earthquake emergency training.

This paper aims to investigate adaptive IVR SGs, focusing mainly on the customization aspect. Research questions include: 1. What are the parameters of a customizable IVR

SG? 2. How do the parameters contribute to the effective customization of an IVR SG? 3. Are these parameters suitable to deliver a customized earthquake training experience? In order to address these questions, we propose a framework for a customizable IVR SG suited to earthquake emergency training. Four essential dimensions and six customization parameters will be outlined to establish the framework. Subsequently, a case study is presented, which includes a prototype developed based on the framework and the validation of its usability, leading to a concluding discussion of challenges and future research directions.

5.2. Background

5.2.1. Adaptation and customization of SGs

An adaptive SG responds better to trainee needs and backgrounds when context-based variations are included (Streicher & Smeddinck, 2016). As a result, trainees receive a unique tailor-made training experience. Thus, adaptive SGs have a wide range of applications. Hardy et al. (2015) introduced adaptive SGs for health sport and physical exercise. In their framework, user profiles and real-time data (e.g., power, movement, heart rate) monitored by sensors are the primary sources of data feeding into training programs to make them more adaptation. In other words, training programs adjust parameters dynamically based on user performance. Kickmeier-Rust and Albert (2010) incorporated adaptive features in an educational game in which secondary school students learned optics. Adaptation took place when the game identified that learners were in need of a different type of pedagogical intervention. A set of trigger mechanisms are defined to deploy the interventions. For instance, learners may face difficulty in solving a problem. Once certain trigger conditions (e.g., no actions, or several failures) are met, a new pedagogical intervention can be applied to assist these learners.

In general, there are two approaches to adaptation, namely system-controlled adaptation and user-controlled adaptation (Orji, Oyibo, & Tondello, 2017). System-controlled adaptation relies on real-time training feedback to assess trainees so that adaptation can be made dynamically. In contrast, user-controlled adaptation provides customization features for trainees to make adjustments by themselves (Orji et al., 2017). Either approach achieves adaptation according to the particular trainee characteristics such as prior knowledge and experiences, preferences, learning progress, or training objectives (Streicher & Smeddinck, 2016). Adaptive training pays careful attention to these aspects to meet training needs and reduce learning gaps.

System-controlled adaptation is an autonomous process that requires minimal to no effort from trainees. It assesses the competence and learning progress of trainees by monitoring their training performance. This form of system-controlled adaptation includes the balance of game difficulty and the deployment of educational interventions (Kickmeier-Rust & Albert, 2010; Streicher & Smeddinck, 2016). In this case, system-controlled adaptation takes place at a micro-level with limited ability to alter the training context or objectives (Kickmeier-Rust & Albert, 2010; Kickmeier-Rust, Augustin, & Albert, 2011; Streicher & Smeddinck, 2016). In contrast, user-controlled adaptation occurs at a macro level offering more flexibility to customize training objectives (Kickmeier-Rust et al., 2011; Orji et al., 2017; Streicher & Smeddinck, 2016). User-controlled adaptation is a manual process that requires configuration from trainees, with guidance from trainers. This process happens before the actual training, tuning parameters in order to shape the training content and experiences. There is a broad range of variations that can be made by trainees with corresponding parameters provided by training programs (Orji et al., 2017; Streicher & Smeddinck, 2016).

In the context of earthquake emergencies, external factors such as trainee type (i.e., the roles trainees will hold in an earthquake event, such as evacuees, floor wardens, or emergency response personnel) play a fundamental role in framing the training. One training program for a specific trainee type can be irrelevant or less effective for another trainee type. For instance, training for children and training for professional personnel (or adults) have to be different in terms of pedagogical strategies, narratives, context, and training objectives. The differences occur at a macro level, which makes user-controlled adaptation a better solution for customization. In this paper, we propose a customizable IVR SG framework based on the user-controlled adaptation approach.

5.3. Customizable IVR SGs for Earthquake Emergency Training

Customization is a manual process that supports training programs allowing adaptability as a key feature (Streicher & Smeddinck, 2016). Adaptability is embodied in parameters that are presented to trainees in a settings' interface (Streicher & Smeddinck, 2016). Such parameters, in turn, can represent the variations in training. De Freitas and Oliver (2006) propose a game-based learning model that includes four fundamental dimensions, namely learner specifics, pedagogy, representation, and context. According to the model, learner specifics cover learner profile, role and competencies; representation refers to the fidelity and interactivity of content; context means the virtual environment where training takes place; pedagogy includes pedagogical approaches (van Staalduinen & de Freitas, 2011). Apart from the game-based learning model, Feng et al. (2018) proposed a framework for developing and implementing IVR SGs for evacuation research. In that framework, the gaming environment covers six elements, which are teaching methods, navigation, hazards simulation, narratives, Non-player Characters (NPCs), and sensory stimulation. Among them, navigation (control mechanisms for navigating in virtual environments), NPCs (to be populated representing other evacuees), and sensory stimulation (additional stimulation outside IVR such as touching and olfactory) are targeting behavioral analysis. These three elements are excluded from the consideration for our customization framework, as the framework aims to deliver an IVR SG with a training purpose. The remaining three elements can be incorporated into the game-based learning model, with teaching methods under the pedagogy dimension, hazards simulation under the context dimension, and narratives under the representation dimension. As a result, we propose six customizable parameters belonging to four dimensions of game-based learning model, which are relevant to IVR SGs-based earthquake emergency training. These parameters make up the framework of customizable IVR SGs for earthqcuake emergencies training, as shown in Figure 5-1. The learner specifics dimension is a root dimension which determines the settings of other dimensions. Storytelling methods are applied to storylines regarding the ways to narrate, and earthquakes and damage are applied to virtual environments to obtain an earthquake simulation. More detailed explanations about this framework are discussed in the following sections.

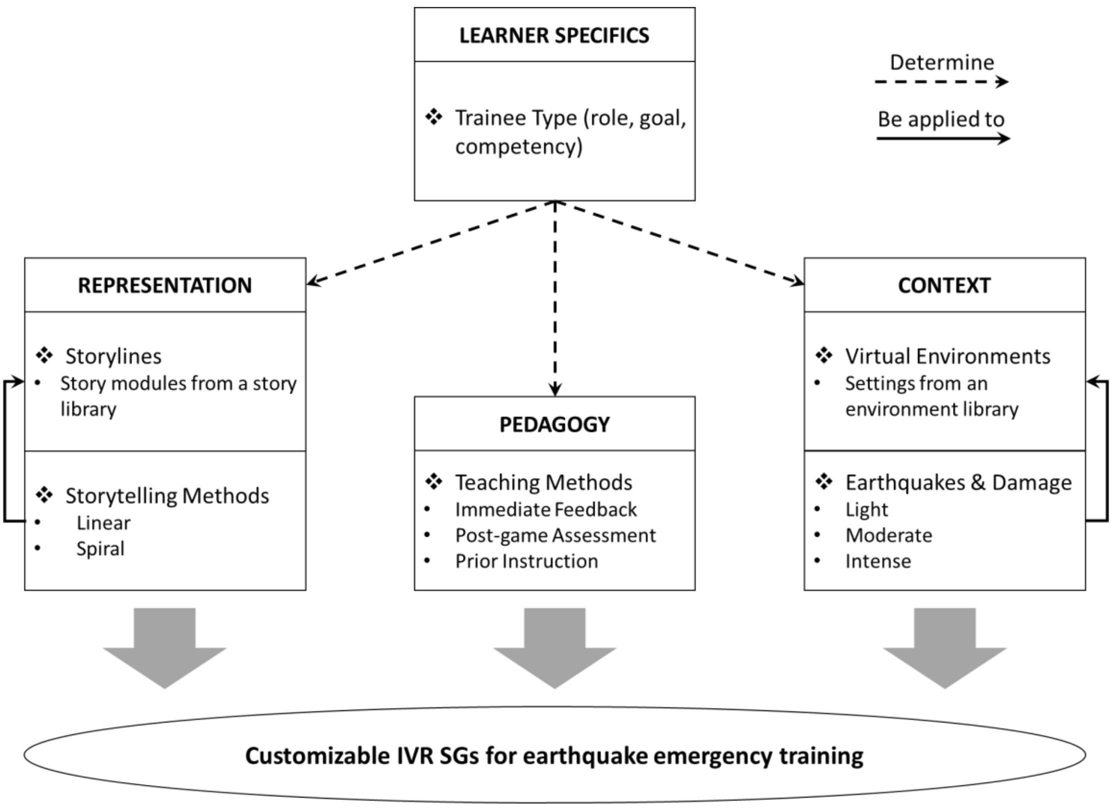


Figure 5-1 The framework of customizable IVR SGs for earthquake emergency training, with six parameters to tune four dimensions.

5.3.1. Trainee type

As part of the learner specifics, trainee type is the heart of user-controlled adaptation. Trainee type is the fundamental parameter of the entire system as it ultimately defines the training objectives. During earthquakes and post-earthquake evacuation, building occupants have different responsibilities and behavioral responses depending on their roles pertaining to the buildings they are in. For instance, in New Zealand, school teachers are required to get students away from dangerous areas, and check students against class lists after earthquakes (Ministry of Education, 2019), while students are instructed to listen to and follow all instructions from adults or the radio (Ministry of Civil Defence and Emergency Management, 2019); for hospitals, staff members have the responsibility to take care of patients and visitors while patients and visitors need to follow general evacuation practice (Auckland District Health Board, 2014; New Zealand Ministry of Civil Defence & Emergency Management, 2015). These recommendations of correct behavioral responses are acknowledged as training objectives to be delivered by IVR SGs (Lovreglio et al., 2018). Relevant learning content is presented to trainees in the form of scenarios that are part of training storylines. As a result, a trainee type determines the storylines of IVR SGs, enabling personalized training content. As well as this, trainees have different competencies and learning styles. Pedagogical strategies, which are influenced by trainees themselves, need to be applied accordingly. Lastly, a broad range of trainees may be familiar with different types of built environments. In this case, trainee type also has an impact on which virtual environment is to be applied to the training in context.

5.3.2. Storylines

According to Göbel et al. (2010), storylines can be deconstructed into story modules, and each story module represents a single scenario. Narratives connect a set of story modules to form a storyline. A storyline consists of a series of story modules where knowledge is conveyed in context (Starks, 2014). For customizable IVR SGs, storylines should be adaptable to the preferences of trainees in order to deliver tailored training objectives. An individual story module can represent a training objective. For instance, if “Help others if you can” is one of the training objectives, then a story module could be a scenario where someone (e.g. a Non-player Character) is trapped and asking for help. In this story module, trainees are taught that they should help others if they can during earthquake evacuation (Ministry of Civil Defence & Emergency Management, 2019). We were inspired by Göbel et al. (2010) to propose a modular storyline model for customizable IVR SGs, as shown in Figure 5-2.

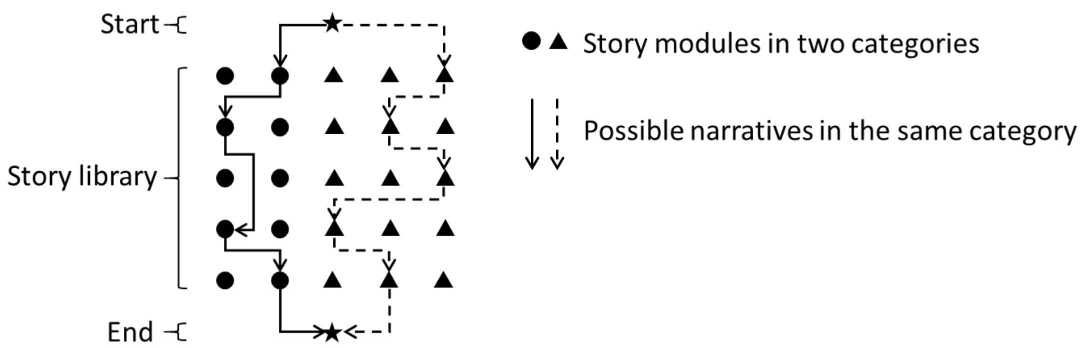


Figure 5-2 Modular storylines. Dots and triangles are story modules differentiated in two categories. Arrow lines represent possible narratives that link story modules in the same category to make a compatible storyline.

A set of story modules are collected as a story library. Following a modular approach, story modules can be picked from the story library and sequenced by narratives to form up a storyline. The modular approach gives the flexibility to generate storylines based on different training objectives, which makes it possible to apply macro adaptation to storylines. Macro adaptation refers to the sequencing of learning events (i.e., which story module to choose next) (Göbel et al., 2010; Kickmeier-Rust et al., 2011). With macro adaptation, trainees can select and skip story modules to make personalized storylines for their training.

However, macro adaptation may face difficulty regarding the control over storylines between trainers and trainees (Kickmeier-Rust et al., 2011). To be specific, some story modules in a story library are not compatible with a certain type of trainees. For instance, a student trainee does not need to go through a story module targeting teachers, in which “check students against clast lists” is the main learning content. Trainees may select story modules that are not suitable to them when they face all the story modules in a story library. A possible solution to the compatibility issue of story modules is to categories story modules. Story modules are tagged and stored in different categories in a story library. Once a trainee type is defined, only the relevant story modules in the same category will be made available from the story library for macro adaptation. As shown in Figure 5-2, dots and triangles are story modules in two categories. Narratives link story modules in the same category to form consistent storylines. Macro adaptation should take place following the rules defined by the trainers and the needs of the trainees (Kickmeier-Rust et al., 2011).

Although storylines are split into individual story modules, the coherence of an entire storyline should be maintained. In the case of earthquake emergency training, the overarching theme of the storylines should be related to specific behavioral responses and safety knowledge about earthquakes and post-earthquake evacuation in a certain setting. Best evacuation practices and government guidelines are often the recommended source to obtain the applicable knowledge of training objectives for storylines (see Table 5-2 as an example) (Lovreglio et al., 2018). Individual story modules should focus on the recommended behavioral responses and present them in the scenarios (Lovreglio et al., 2018). While the main focus of storylines is to present the applicable knowledge, the transfer of knowledge (i.e., teaching) is achieved by the pedagogies applied (see Section 5.3.4).

5.3.3. Storytelling methods

IVR SGs need to maintain trainees in a state of flow. Flow is a state when individuals get highly engaged in activities (Czikszentmihalyi, 1990). In this state, individuals perceive that performing an activity in the IVR SG brings them enjoyment and rewards; therefore, they are able to reach a high-level functional capacity to manage the activity (Admiraal, Huizenga, Akkerman, & ten Dam, 2011; Nakamura & Csikszentmihalyi, 2014). A flow state is likely to occur when the challenges of activities in IVR SG are in balance with individuals' skills (Admiraal et al., 2011). If challenges are beyond the capability of trainees, the challenges can become overwhelming and cause anxiety. In turn, if challenges fail to engage trainees, then trainees can lose attention and get bored (Admiraal et al., 2011). In the context of IVR SGs, challenges that are embedded in story modules are fixed, based on training objectives. In order to reach a flow state, trainees need to achieve small manageable steps successfully and obtain large and complex training objectives eventually (Dunwell et al., 2011). Different storytelling methods can be applied to achieve this.

Storytelling methods, also known as narratives, refer to how storylines (content) are discoursed and presented to audiences (Chatman, 1980). Narratives are crucial to communicating learning content to the trainees of IVR SGs (Lim et al., 2014). Narratives can either be explicit or implicit. An explicit narrative means storylines are clearly structured and outlined to trainees, while an implicit narrative refers to a more open situation where trainees can explore without storylines being structurally presented (Ferguson, van den Broek, & van Oostendorp, 2020). In a recent study, Ferguson et al. (2020) note that explicit narratives, as compared to implicit narratives, deliver better performance in terms of transferring factual knowledge. One possible way to make an explicit narrative is to use an action-driven method, which drives storylines based on the actions taken by trainees (Feng et al., 2018). With this method, each story module explicitly contains one small and manageable task for trainees to solve. Trainees have to take action in each story module so that they can make progress through storylines following the defined narratives in a linear way. Such linear action-driven narratives help trainees focus on one small problem at a time, which enables the training experience to flow for trainees.

In the case when the level of difficulty might be beyond the capability of trainees, spiral narratives can be an alternative for trainees to choose to rebalance the difficulty and capability, as shown in Figure 5-3. Specifically, if trainees take a wrong action in a story module, storylines can be paused. Feedback is given to the trainees, which can be in various formats, including visual effects (e.g., a red cross for a wrong action), sound effects (e.g., a wrong action sound), and explanatory texts. Then the same story module is repeated. Storylines will not progress unless the trainee takes the correct action in the story module. In this way, trainees can make several attempts, which can lessen anxiety and keep trainees in a state of flow. After their initial attempts, trainees can reflect on their actions and correct their mistakes based on given feedback. Such reflection and revision help trainees avoid the same mistakes being made again, which is essential when facing actual situations in the future (Scoresby & Shelton, 2014).

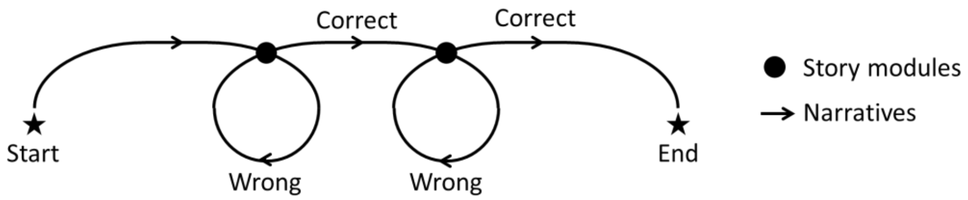


Figure 5-3 Spiral action-driven narrative

The term “spiral” originated from the concept of “spiral learning experience” (Henton, 1996). As shown in Figure 5-4, experiential learning develops perspectives and skills in a spiral way, including several iterative phases, such as experience, reflect, generalize, and transfer. Similarly, regarding spiral narratives, trainees are prompted to progress their learning in a spiral way, not in a loop process.

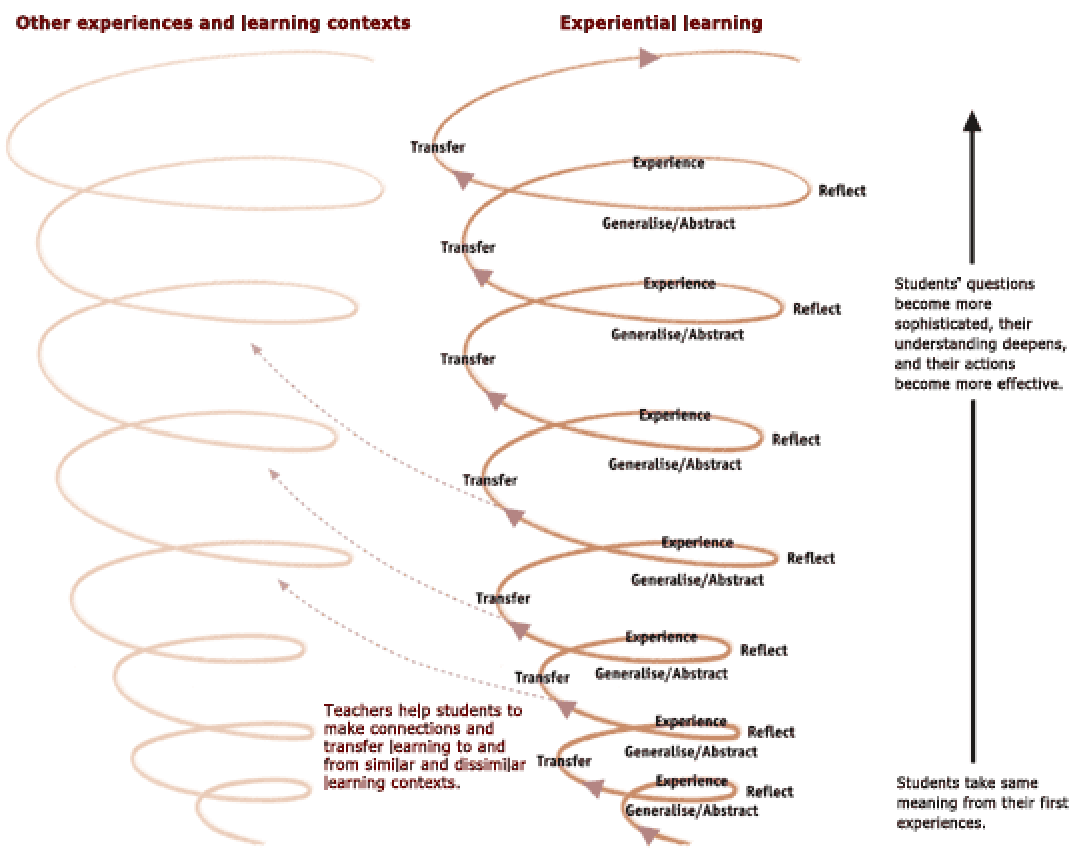


Figure 5-4 Spiral learning experience (Henton, 1996)

5.3.4. Teaching methods

Teaching methods, in the context of IVR SGs, are the pedagogies presented to trainees to achieve the training objectives. In general, evacuation training conducted by IVR SGs are simulations based on experiential models. Trainees are placed in immersive 3D environments where they experience emergencies and take action following the best evacuation practice. Learning takes place during the experience. Based on problem-based learning (Forcael et al., 2014), Kiili (2007) proposes a problem-based gaming framework, by which trainees learn about a subject through the experience of solving tasks. Tasks are embedded in simulations to facilitate problem-based learning. Tasks are generated from the training objectives that are used to make story modules. Each story module contains a problem for trainees to solve in order to achieve a training objective. For instance, when an earthquake strikes, trainees are required to protect themselves, in which case, “Drop, Cover and Hold" would be the correct solution to this problem (Ministry of Civil Defence & Emergency Management, 2019).

One important phase of problem-based gaming is the reflection phase, in which trainees revisit and evaluate their experiences, and learn from them (Kiili, 2007). Different ways to foster reflective thinking have been used in previous studies of IVR SGs, including immediate feedback, post-game assessment, and prior instruction (Feng et al., 2018). Immediate feedback gives in-game information right after the completion of a task. Post-game assessment, which can be categorized as delayed feedback, provides a summary and explanation after the completion of the entire training storyline. These two forms of feedback carry different sets of information and are presented in different timing sequences. The effectiveness of these two forms of feedback is controversial. Schroth (1992) notes that delayed feedback may be more effective for concept-formation task learning, while immediate feedback may enhance procedural learning. Metcalfe et al. (2009) found that delayed feedback was more effective than immediate feedback in the context of vocabulary learning. Nevertheless, it is argued that the effectiveness of feedback relies on the nature of tasks and the capability of learners instead of the timing (Mathan & Koedinger, 2002). Bearing this in mind, we provide both immediate and delayed feedback in our customizable IVR SG framework for trainees to choose from.

Another approach to facilitate reflection is using prior instruction, where instructions are given to trainees prior to a task to be solved. It is argued that clear instructions have a positive impact on learning motivation, which is beneficial to promote reflective thinking (Erhel & Jamet, 2013; R. E. Mayer & Johnson, 2010). This approach has also been suggested when trainees may lack confidence or sufficient capability to complete a task (Jarvis & de Freitas, 2009). This view is supported by S. Smith and Ericson (2009) who applied instructions to help young participants (7 to 11 years old) focus on training instead of being anxious when facing hazards in a study using IVR SG for fire safety training. In the case where the capability of trainees might be below the difficulties of given challenges, prior instruction may be useful to achieve a flow state by reducing difficulties. With system-controlled adaptation, systems are possible to assess the performance of trainees and alter the pedagogical approaches accordingly (Orji et al., 2017). As for user-controlled adaptation, trainees make such decisions. This approach to teaching is included in our framework as we aim to reach a large audience across all ages and backgrounds.

5.3.5. Virtual environments

In the present study, a virtual environment is a place where trainees are exposed to earthquakes, take actions and acquire safety knowledge. Hansman (2001) argues that a real-world context makes the best learning environment. This point is supported by Yu et al. (2015), who suggest that context-based learning effectively improves problem-solving skills. Therefore, a virtual environment is recommended to represent real-world settings covering various building types, such as hospitals, schools, office buildings, and residential buildings. A library storing multiple settings can deploy suitable virtual environments to trainees according to their preferences.

One possible solution to represent built environments in IVR SGs is Building Information Modelling (BIM) (Rüppel & Schatz, 2011). BIM models carry standalone building objects with geometric and material information, which makes it feasible to model earthquakes and the damage to buildings in virtual environments (Feng et al., 2018). Based on site surveys, including blueprints, drafts, and photos, BIM models permit detailed reconstruction of built environments close to reality (Feng et al., 2018; Lovreglio et al., 2018). In order to provide sufficient context for trainees, BIM models need to represent typical layouts and elements that can help trainees recognize relevant building types. For instance, school environments can be created to consist of classrooms and teaching equipment.

The general workflow of BIM solutions is to import BIM models to game engines for IVR SGs development (Bille, Smith, Maund, & Brewer, 2014). After importing to a game engine, different algorithms can be applied to create shaking experiences, and additional layers of objects and information can be added to simulate hazards and damage (Feng et al., 2018; Lovreglio et al., 2018). In general, there are two algorithms to simulate shaking experiences. One is to shake and manipulate individual objects in IVR. Another is to shake the virtual camera within the IVR environment, plus a few visual effects (e.g., falling dust, cracks, sparks, flying or falling objects). The first approach is able to capture more accurate the earthquake's shaking dynamics, but it involves more development efforts and computational resources. The second one is more efficient and requires less computational resources. In addition, the later approach is as effective as the former approach to obtain the expected pedagogical and behavioural outcomes. This study applied the second approach. Trainees had reported a high level of realism (see Figure 5-12) and increased their knowledge and self-efficacy significantly (see Figure 6-4 and Figure 7-5).

5.3.6. Earthquakes and damage

Earthquakes, which can continue over several months or years, are a sequence of seismic events that can be staged into foreshocks, mainshocks, and aftershocks (Jones, 1994). The complete process and timeframe of earthquakes are not compatible with a game (training) experience (or other typical training methods). Therefore, earthquakes in simulations are an abstraction of actual situations with a number of simplifications, especially for experiential purposes (Lovreglio et al., 2018).

In general, two approaches have been applied to simulate earthquakes, namely quantitative and qualitative (Lovreglio et al., 2017). A quantitative approach performs accurate calculations that heavily rely on complex mathematical models and assumptions. The purpose of a quantitative approach is to predict the actual performance of objects and appreciate the physical relationship between them in an earthquake event (Porter, Kennedy, & Bachman, 2007). The process of a quantitative approach requires a large number of computational resources, and its outcome can be overwhelming for training simulation. Instead, a qualitative approach has been suggested to be more suitable for IVR SG scenarios (Lovreglio et al., 2018). A qualitative approach mimics earthquakes and damage according to qualitative descriptions sourced from video footage and photos of actual earthquakes (Yang & Wu, 2013). An accurate physical appraisal is not intended, which gives more flexibility to develop training programs. In a recent study of IVR SGs, users reported a high level of realism for the earthquake and damage simulation that was developed based on a qualitative approach (Lovreglio et al., 2018). A qualitative approach as a more feasible strategy is able to build a credible training simulation and deliver a reliable user experience.

Earthquakes have different levels of intensity that result in different levels of damage to buildings. Therefore, building occupants should have specific behavioral responses to earthquakes with different magnitude or intensity. For IVR SGs, the differentiation between earthquake intensity allows trainees to face different levels of challenges. Hence, earthquake intensity is another parameter to be considered in our framework. Subjective intensity is one of the forms used to describe earthquake impacts.

Subjective intensity is a qualitative measure describing the observable effects on people, buildings, and objects in the event of an earthquake (BRANZ seismic resilience, 2019). As such, Geonet (New Zealand's national geophysical monitoring system) provides a detailed description of the New Zealand Modified Mercalli Intensity (MMI) Scale, which has twelve levels with 1 representing the weakest and 12 representing the most destructive (GeoNet, 2019). A qualitative approach can be applied to build earthquake simulations with different levels of damage for different levels of intensity based on the Geonet database. For instance, "walls may creak, and glassware, crockery, doors or windows rattle" in a light earthquake, while "furniture and appliances are shifted, substantial damage to fragile or unsecured objects" in a more severe earthquake (GeoNet, 2019).

5.3.7. Summary

The training program is the outcome of the customization process. In a nutshell, this represents a four-dimensional hierarchy containing information that is layered atop one another to build up the customized training program (see Figure 5-5).

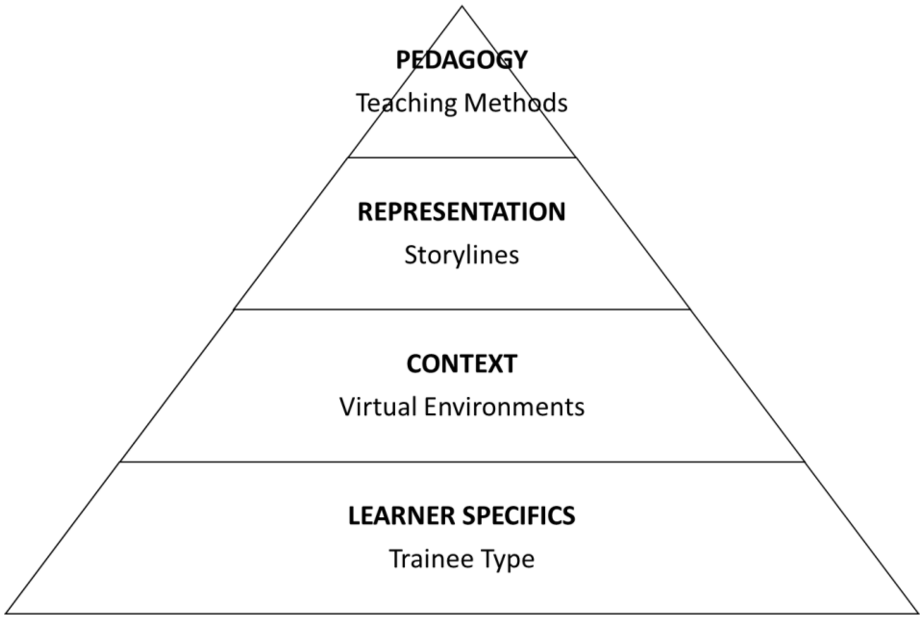


Figure 5-5 Information layers for customized training programs

The base layer is learner specifics which are defined as the tuning of the training program. The context is built upon learner specifics, mapping out virtual environments in which the training takes place. The next layer, representation, is inserted into the context layer presenting training content as storylines. The top layer, pedagogy, is where teaching strategies are deployed to educate trainees. With this layered structure, each dimension is independent, in which case, the information within a dimension is interchangeable without affecting the other dimensions. Trainers benefit from this by having the possibility to refine a single dimension without having to worry about the other dimensions.

A typical process of creating a customized training environment involves trainees in two phases, with phase one being the customization process and phase two being the training itself (see Figure 5-6). In phase one, parameters can be presented in an interface, where trainees make selections as user-controlled adaptation. In the case where trainees are not familiar with the presented parameters, the image-based or text-based explanations of the parameters can come along with the parameters. Trainers can also provide guidance to trainees. In phase two, trainees go through their customized training program.

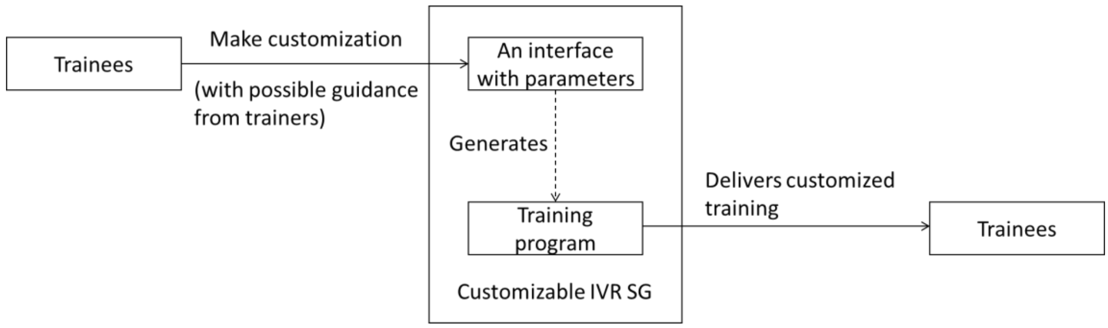


Figure 5-6 The procedure of the implementation of a customizable IVR SG

The parameters of the customization framework of IVR SGs (see Figure 5-1) are discussed in previous sections. Table 5-1 summaries the manipulation of these parameters in phase one, from a trainee's perspective. The customization takes place as in the hierarchy shown in Figure 5-5, where trainees customize learner specifics first, then context, followed by representation, and last pedagogy.

Table 5-1 The manipulation of the parameters by trainees

Dimensions	Parameters	Manipulation by trainees
Learner specifics	Trainee type	Trainees make a selection based on their personal profiles and actual roles in the event of an earthquake.
Context	Virtual environments	Trainees make a selection from a library storing BIM-based virtual environments.
	Earthquakes and damage	Trainees make a selection for

		earthquake intensity, which simulates earthquakes and damage based on qualitative descriptions (Modified Mercalli scale).
Representation	Storylines	Trainees select story modules in the same category from a story library to form a storyline which is compatible with trainee type.
	Storytelling methods	Trainees make a selection from linear and spiral narratives based on their capability
Pedagogy	Teaching methods	Trainees make a selection from immediate feedback, post-game assessment and prior instruction based on their preferences

5.4. Case study

A prototype was developed by implementing the proposed framework. Pilot trials were conducted to validate the framework. This section presents the prototype as well as the participants, measures and results of the pilot trials.

5.4.1. Materials

The first step of the prototype development was to identify target trainees. Two types of trainee were targeted for our case study, and these were children (junior secondary school students) and adults (university students and staff). For these two trainee types, training objectives were identified accordingly, as shown in Table 5-2 (Ministry of Civil

Defence and Emergency Management, 2019; Ministry of Education, 2019; New Zealand Ministry of Civil Defence & Emergency Management, 2015). Story modules were created and categorized based on each training objective so that a story library was established.

Table 5-2 Training objectives and story modules with categories

Training objectives	Story modules	Categories
Drop, Cover and Hold in an earthquake	An earthquake strikes the building trainees are in. Trainees need to decide how to respond.	Children, Adults
Do not attempt to run outside during the shaking	Trainees can decide whether to run during the earthquake.	Children, Adults
Listen to and follow all instructions from adults or the radio	After the shaking, trainees can decide whether to follow a teacher to evacuate	Children
Stay calm. If you can, help others who may need it	A person (Non-player Character) is trapped under a table. Trainees can decide whether to offer help.	Children, Adults
Watch out for possible hazards	Massive broken glass is on the evacuation route. Trainees need to decide whether to use an alternative route.	Children, Adults
Remember there may be some aftershocks	An aftershock comes when trainees are evacuating. Trainees need to decide how to respond.	Children, Adults
Take personal items when you	Before evacuation, trainees need to	Adults

evacuate	decide whether to take their wallet, cellphone, and laptop.	
Do not use a lift when you evacuate	Trainees need to decide which evacuation route to use.	Adults
Mitigate hazards if you can	An electrical appliance is emitting sparks. Trainees need to decide whether to turn off the power supply.	Adults

Regarding virtual environments, two settings were developed to be stored in an environment library, including a junior secondary school setting and an office building setting (see Figure 5-7). Variations of earthquake simulation and damage were applied to these settings, which consisted of three levels: light (MMI 4), moderate (MMI 5), and intense (MMI 6) (GeoNet, 2019). Such variations were achieved by changing the number of falling and breaking objects (i.e., visual effects) as well as altering soundtracks (i.e., sound effects).

(a1) The virtual school environment before an earthquake

(a2) The virtual school environment after an earthquake (MMI 6)

(b1) The virtual office environment before an earthquake

(b2) The virtual office environment after an earthquake (MMI 6)

Figure 5-7 Virtual environments and earthquake damage: (a) The school setting, (b) The office setting

Upon the completion of a virtual environment, story modules were deployed in different locations within the environment. A waypoints system as a navigation method was adopted to line up the story modules. Waypoints are sets of stopping points or direction-changing points on a navigation route (Ragavan et al., 2011), with each story module representing one waypoint. A navigation route was developed connecting waypoints. The movement of trainees was limited to the navigation route, in which case they only moved their vision from one waypoint to another (in our case, from one story module to another). For the linear narrative, trainees were led from one waypoint to the next disregarding what they did at each waypoint (i.e., story module). While in the case of the spiral narrative, trainees stayed and revisited a waypoint (i.e., story module) if they did something wrong, and they could only make progress to the next one after they solved the challenges set in the actual waypoint correctly. Figure 5-8 shows an example of a waypoints system with the integration of story modules and a virtual environment.

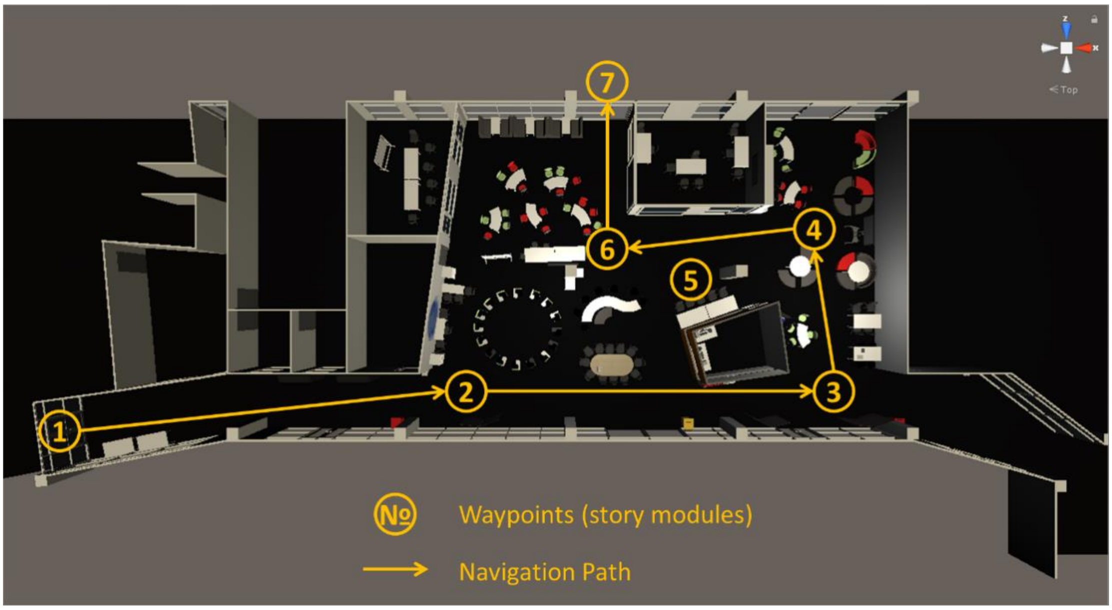
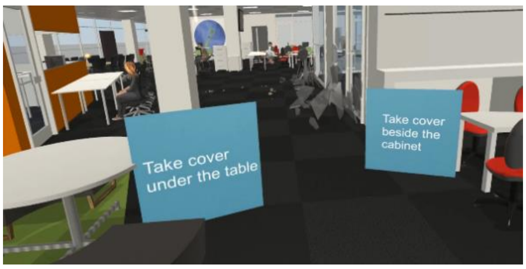


Figure 5-8 An example of a waypoints system deployed in a school setting where waypoint 5 (story module 5) is not included in the storyline.

Finally, the pedagogy dimension was added to the prototype. As challenges were presented to trainees to solve in each story module, teaching methods with three options were applied to fulfil pedagogical objectives (see Figure 5-9). For immediate feedback, comments were presented to trainees right after they completed a challenge. In terms of post-game assessment, trainees received a final report and an explanatory video upon the completion of the entire training. Regarding prior instruction, text-based guidance was provided to trainees prior to each challenge.



(a1) Trainees experienced an event and tried to select an action choice in response to the challenge

(a2) Immediate feedback was given to trainees after solving the challenge

(b1) At the end of the storyline, a checklist is presented to trainees as post-game assessment

(b2) Trainees revisited their previous training experience through a playback video

(c1) Prior instruction was given to trainees before a challenge

(c2) After receiving instructions, trainees experienced an event and tried to select an action in response to the challenge

Figure 5-9 The instructional approaches of the IVR SG training system: (a) Immediate feedback, (b) Post-game assessment, (c) Prior instruction

At this stage, four dimensions were able to be customized in the prototype, as discussed in Section 5.3. An interactive user interface with adjustable parameters and the explanation of each parameter was presented at the beginning of the IVR SG to trainees in order to complete the customization process before actual training (see Figure 5-10).

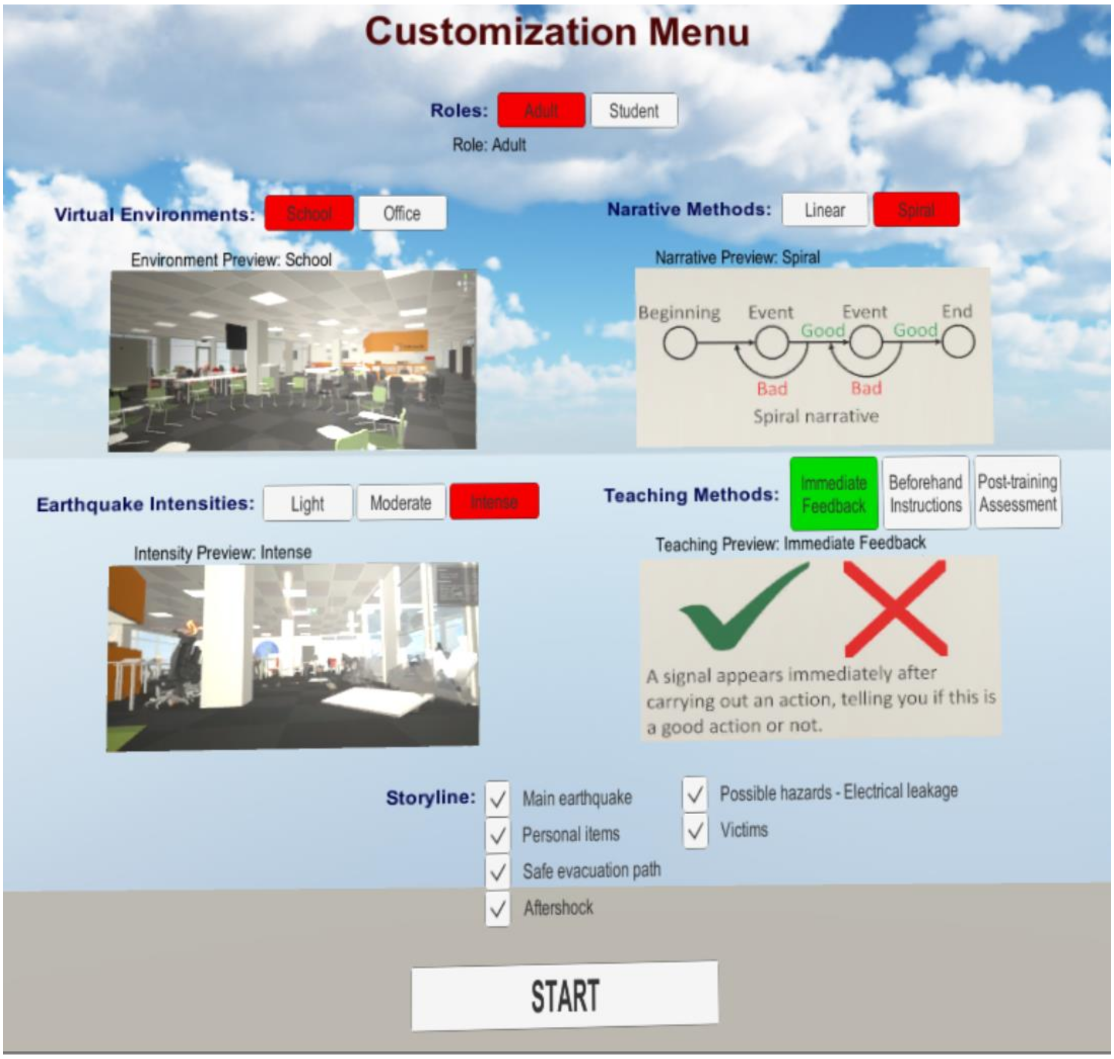


Figure 5-10 An interface with parameters enables the customization process

The prototype was developed and implemented using Unity 2018.2.14f1 (www.unity.com), which is a game engine for video game development. IVR was enabled by Oculus Rift which comes with a head-mounted display (HMD) IVR solution. Oculus Rift enables a 110-degree field of view both horizontally and vertically, with 1080x1200 resolution per eye. The interaction with virtual objects in the IVR environment was achieved by Oculus touch controllers, which are the accessories of Oculus Rift. Participants could point at a virtual object with a green laser beam using the controller and select an option by pressing any button in IVR. With the interface, as shown in Figure 5-10, participants could use the Oculus touch controller to point at and select options for each parameter. Participants could use the same interaction mechanism to choose actions when they are facing challenges to solve during training (see Figure 5-9). In terms of navigation, the movement of participants in virtual environments followed a predefined path using a waypoints system (see Figure 7). The way of interaction and navigation assisted participants in focusing on training tasks, with minimum distractions on gaming controls and the exploration of virtual environments.

5.4.2. Participants

A total of 191 participants were recruited by posters, email, and referrals to test the prototype. Among them, 125 participants were children (junior secondary school students, female: 46, male: 79) and 66 participants were adults (university students and staff, female: 34, male: 32). Children were aged from 11 to 15 years (M = 12.4, SD = 1.08) while adults were aged from 18 to 53 years (M = 27.7, SD = 8.40). These two groups of participants resulted in two trainee types for the training prototype, with one type being children who were familiar with school settings and the other being adults who were used to office settings.

Since the measures of the prototype were mainly focused on usability and user experience (as discussed in Section 0), participants were questioned about their familiarity with IVR and video games, as shown in Table 5-3 and Table 5-4. Kruskal-Wallis H tests confirmed that no significant difference existed between these two groups for IVR experiences ($p = 0.059$), while a significant difference for video games ($p < 0.001$). Children might adapt more easily to game-based concepts.

Table 5-3 Experiences with IVR in the form of HMD

	Children (125)		Adults (66)	
Yes	82	65.5%	34	51.5%
No	36	28.8%	32	48.5%
Unsure	7	5.6%	0	0.0%

Table 5-4 Frequency of playing video games on phones, computers, tablets, consoles, etc.

Frequency	Children (125)		Adults (66)	
Never	2	1.6%	8	12.1%
Less than once a year	2	1.6%	14	21.2%
At least once a year	1	0.8%	10	15.2%
At least once a month	7	5.6%	13	19.7%
At least once a week	28	22.4%	10	15.2%
Several days a week	38	30.4%	9	13.6%
Every day	47	37.6%	2	3.0%

The adequacy of the sample size was assessed. According to meta-analysis studies on Virtual Reality Exposure Therapy (Carl et al., 2019; Morina et al., 2015), the effect size of IVR training for the entire population was medium to large. In this case, Cohen's d (i.e., the effect size) was estimated to be 0.5 (J. Cohen, 2013). A power analysis using the statistical software G*Power 3.1 indicated that a total sample of 35 people would be needed to detect a medium effect (d = 0.5) with 80% power using a t-test. A total sample of 56 people would be needed to detect a medium effect (d = 0.5) with 95% power using a t-test. In addition, Gall et al. (2006) suggest a sample size of larger than 15 cases for causal-comparative and experimental methodologies. The sample sizes of other IVR studies with similar methodologies fall in the range from 16 to 59 for each sample group (Li et al., 2017; Lin et al., 2019; Lin et al., 2020). In this study, it is argued that the sample size is adequate for the experiments undertaken; however, the author is still aware that large scale experiments with random sampling are necessary to generalize the findings to a larger population (L. Cohen et al., 2013).

5.4.3. Measures

This study aims to propose a customization framework and validate it through a prototype. We followed a prototyping development process, in which case its evaluation focuses mainly on the usability of the prototype (Lovreglio et al., 2018). In our case, good usability and an enjoyable user's experience also play an important role in user-controlled adaptation (Streicher & Smeddinck, 2016). This includes the interaction with the customization interface as well as the outcome of the customization (i.e., customized training experience). Therefore, the System Usability Scale (SUS) was used to develop the statements of our questionnaire (Brooke, 1996). While SUS evaluates the overall situation of a system, we shaped the statements to pay attention to five aspects of our customizable IVR SG, as shown in Table 5-5. Statements were reworded in plain language in order to be understandable for children.

Table 5-5 Statements assessed in the questionnaire

Aspect	Statements
Customization process	The customization menu was easy to understand. The instructions on the menu were easy to follow.
Virtual environments	The building environment was realistic. The IVR experience was realistic.

Storylines	The training storyline helped me learn.
	It was easy for me to understand the learning content.
Teaching methods	It was easy for me to learn about what to do during and after earthquakes.
	It was easy for me to remember what I have learned.
Engagement	It was easy for me to concentrate on my learning.
	It was easy for me to stay focused on the task.
	I felt the training was fun.

For the adult group, a 7-point Likert scale (-3 = strongly disagree, +3 = strongly agree) was adopted for rating their levels of agreement to these statements. For the child group, a different response format was used. Number-based Likert scales have been found to be unreliable with children as they may have trouble to reflect their perspectives with numbers (Mellor & Moore, 2013). Instead, a response format with words describing the frequency of thoughts has been suggested (Mellor & Moore, 2013). In this case, we applied an explicit text-based scale for the child group, which included four levels of descriptions: never true, hardly ever true, sometimes true, and always true. Accordingly, these four levels were coded from 1 to 4 for data analysis, where 1 represented “never true”, and 4 represented “always true”. The reason to keep a 7-point Likert scale instead of the four-level text-based scale for adults was that 7-point Likert could provide more points of discrimination, which could potentially capture more accurate responses and increase the variance in measures in this age group.

5.4.4. Procedures

Pilot trials were conducted at two different locations, Ormiston Junior College (a junior secondary school) in Auckland and The University of Auckland. Figure 5-11 shows the pipeline of pilot trial procedures. Prior to trials, adult participants received and signed a consent form by which they gave consent to take the trial and for their data to be collected anonymously. Children participants signed a consent form, as well as their teachers and parents. The experiment was approved by The University of Auckland Human Participants Ethics Committee, and the board of trustees of Ormiston Junior College. Participants could withdraw from the trial at any time without any reason. Upon arrival, participants answered a questionnaire about their demographic information as well as their prior experiences with IVR and video games. Then, participants were led through an introduction about using IVR, including controls and health and safety. After that, participants put on the HMD and began the IVR session.

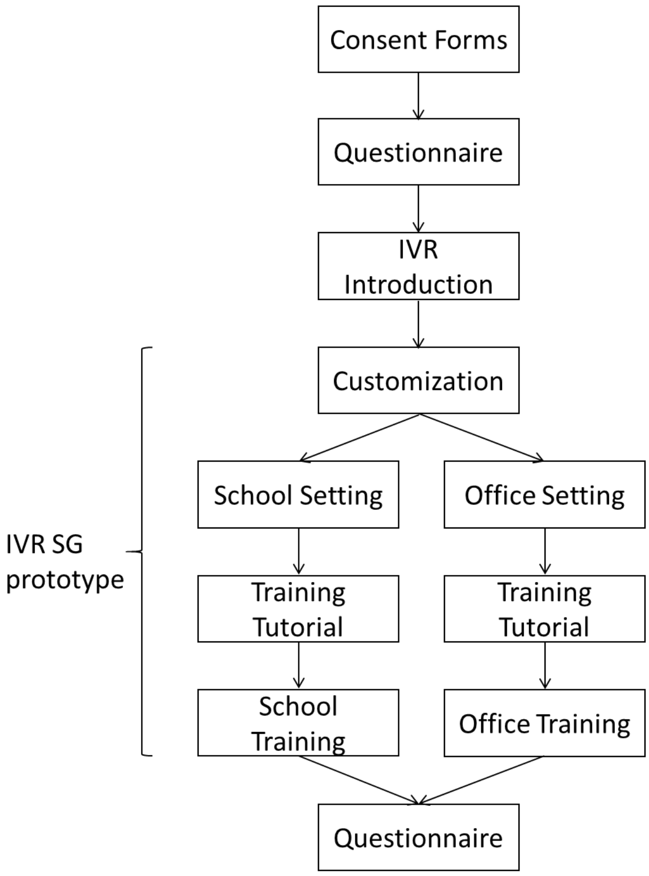


Figure 5-11 Pilot trials procedure

The IVR session included two phases, which were a customization phase and a training phase, as we discussed in 5.3.7. Participants were instructed about making customization through the interface (see Table 5-1 and Figure 5-8). Children were instructed to build a school setting by selecting the “Students” and “School” parameters, while adults were instructed to build an office setting by selecting the “Adults” and “Office” parameters. In order to cover every setting in other parameters equally (e.g., three teaching methods, two storytelling methods), participants who were randomly assigned to different testing groups were instructed to select the parameters based on their testing groups. They had to complete the customization with all the parameters so that to progress to the next phase. Then, a tutorial about interaction and navigation for training took place before the training session. Finally, the corresponding training session started.

Once the entire IVR session was complete, participants answered a questionnaire about the ten statements specified in Table 5-5. Finally, once participants responded to the questionnaire, they were thanked for their participation.

5.4.5. Results

Since two measurement scales were used for the two types of participants, results are reported in two groups. In the case where more than one statement was used to assess an aspect, multiple responses were averaged. For each aspect, the mean score was obtained by averaging the scores from each participant in the same group, with a standard deviation calculated. Cronbach’s alpha was calculated to assess the internal consistency for each aspect. Table 5-6 shows a summary of these results.

Table 5-6 Questionnaire results for two types of participants

Aspect	Children (1 to 4)	Adults (-3 to +3)
Customization process	M = 3.64, SD = 0.558 Cronbach's alpha = 0.806	M = 2.45, SD = 0.680 Cronbach's alpha = 0.779
Virtual environments	M = 3.20, SD = 0.683 Cronbach's alpha = 0.719	M = 2.17, SD = 0.866 Cronbach's alpha = 0.921
Storylines	M = 3.50, SD = 0.549 Cronbach's alpha = 0.664	M = 2.36, SD = 0.727 Cronbach's alpha = 0.806
Teaching methods	M = 3.53, SD = 0.563 Cronbach's alpha = 0.605	M = 2.26, SD = 0.795 Cronbach's alpha = 0.817
Engagement	M = 3.53, SD = 0.524 Cronbach's alpha = 0.701	M = 2.35, SD = 0.667 Cronbach's alpha = 0.737

Note that the Cronbach's alpha of storylines for the child group was below 0.7, which suggests that the statements measuring this aspect are not strongly correlated. However, two storytelling methods (linear and spiral) were tested with the child group. The differentiation of perspectives was not unexpected. It was the same that a Cronbach's alpha of the teaching methods for the child group was less than 0.7, with three methods (immediate feedback, post-game assessment, and prior instruction) tested.

The distributions of participants' responses are illustrated in box plots, as shown in Figure 5-12.

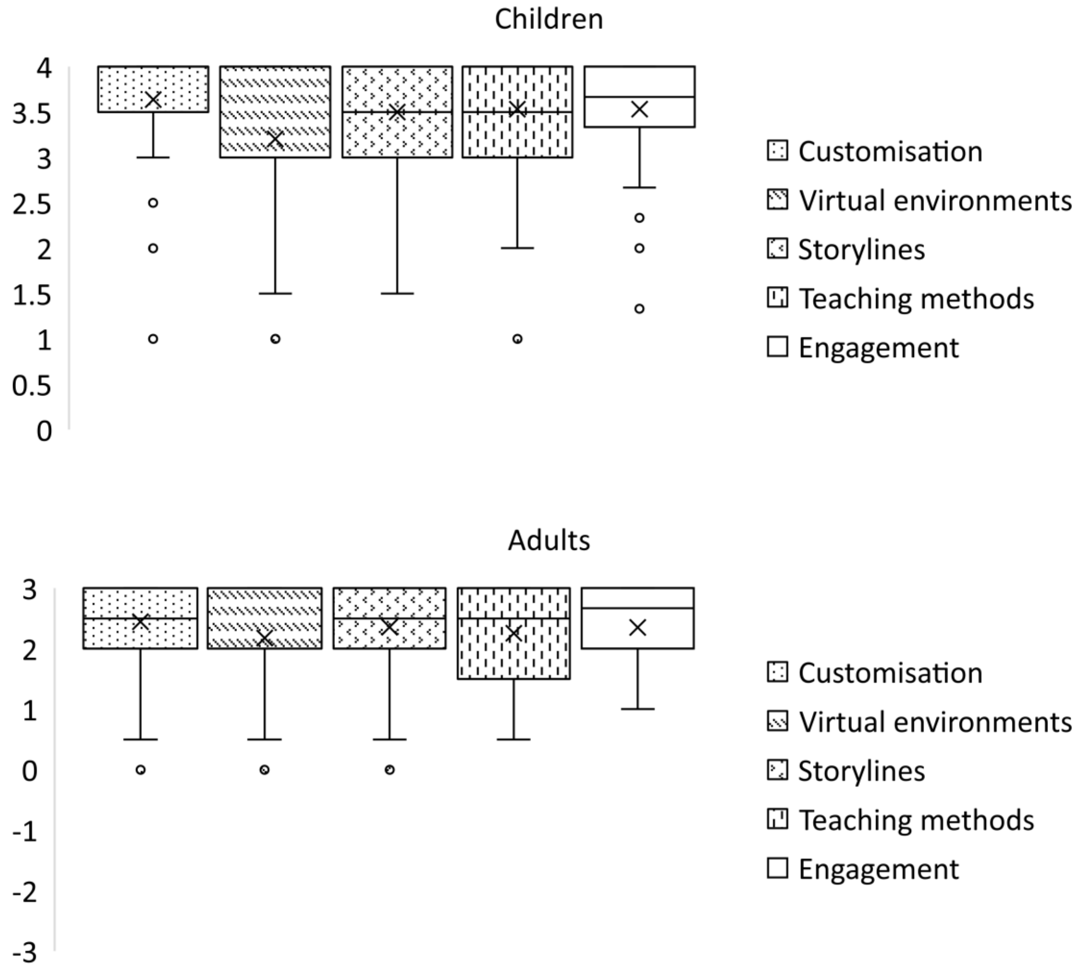


Figure 5-12 Participants responses reported in box plots for children and adults

For each assessed aspect, the average scores were positive that lay in between “agree” to “strongly agree” (“sometimes true” to “always true” for children). Both the child and adult groups reported a high level of acceptance towards the customization process and the outcome of customization (i.e., the customized training experience).

As we measured in pre-test questionnaires, the child group was more familiar with video games than the adult group (no difference between IVR experiences). It is possible that children might be more able to adapt to the customizable IVR SG framework. Therefore, we examined whether there was a statistical difference between the five usability aspects reported by children and adults. Considering that the scores were reported on different scales, the scores were standardized using z-scores (Milligan & Cooper, 1988). Following that, Mann-Whitney U tests were conducted revealing no statistical differences related to the type of trainees for virtual environments ($p = 0.894$), teaching methods ($p = 0.179$), and engagement ($p = 0.265$). Statistical differences were found for customization ($p = 0.018$) and storylines ($p = 0.050$). Adults felt easier to understand the customization than children, and children perceived that learning through storylines was easier than adults. The inconsistency of the results when comparing groups seem to suggest that the familiarity with IVR and video games do not make a significant impact on the user's experience, in terms of the customization process and the customized training.

5.5. Discussion

This paper proposes a framework for customizable IVR SGs in the context of earthquake emergencies training. A prototype was developed based on the principles of the framework. Pilot trials were carried out to evaluate the prototype. Results indicate that the customization process was easy to understand, and the customization was suitable to deliver a customized training experience.

Participants conducted the customization process through an interface. Adjustable parameters and brief introductions were given to participants in the interface. A total of six parameters mapped to four essential dimensions of game-based learning were given, as discussed in 5.3. The ease of use of the customization process was recognized by participants, with the score of 3.64 from children (possible scores from 1 to 4) and the score of 2.45 from adults (possible scores from -3 to 3). This finding was expected as we intentionally offered only a small number of parameters for tuning, which could reduce the complexity of the process. Efficiency and effectiveness are the key factors of quality user experience with the interactions of systems (Hartson & Pyla, 2012). In order to achieve this, monolithic one-dimensional settings have been widely adopted by the game industry (Streicher & Smeddinck, 2016). For instance, game difficulty settings such as "easy" or "hard" reflect multiple game variables such as the number of opponents, punishment mechanisms, and available resources. This kind of trade-off enables the IVR content and challenges to be matched with players with varied capabilities and demands. In our case, one parameter has a direct influence on the corresponding dimension. For instance, for earthquakes and damage, participants only made selections from "light", "moderate", and "intense", with images provided describing each option. Monolithic variations were deployed to a given virtual environment, including sound and visual effects in order to provide three levels of earthquake simulation. Participants were not required to manipulate the details in-depth so that they could pay more attention to the training part.

The advantage of monolithic settings is that less effort is required from trainees to complete the customization process. However, a drawback to this is that the flexibility of the customization is limited. As for user-controlled adaptation, users (trainees) are responsible for making adaptations (Orji et al., 2017). If trainees are facing an open world (i.e., a virtual world where trainees can explore and approach objectives freely) that is full of possibilities to make the customization such as Minecraft (Lane & Yi, 2017), much customization effort from trainees is required to get a training program fitting the training purpose. Furthermore, trainees can be easily distracted from the main training purpose by manipulating various parameters (Orji et al., 2017). Thus, there are always some baselines set by trainers to guide the customization, as the main principle of macro adaptation suggested by Kickmeier-Rust et al. (2011).

Customization, therefore, is an outcome of the combination of trainees' input and trainer's settings, reaching a balance between customization flexibility and main training purpose.

After the customization process, the next session that the trainees went through was the training program. The training program was assessed by our pilot trials. The results for virtual environments (Children: M = 3.20, SD = 0.683; Adults: M = 2.17, SD = 0866), storylines (Children: M = 3.50, SD = 0.549; Adults: M = 2.36, SD = 0.727), and teaching methods (Children: M = 3.53, SD = 0.563; Adults: M = 2.26, SD = 0.795) indicate that the customization was suitable to generate a well-accepted customized training program, regardless of which type of participant. Learner specifics were not measured as they were determined by trainee type, in our case, the children and adults.

One possible explanation for the expected outcomes of the customization process relates to how the layered structure could be modularization. Modularized content plays an important role to effectively interchange information within a dimension (Streicher & Smeddinck, 2016). We proposed modular storylines (see Figure 5-2). In our case, as illustrated in Figure 5-13, training objectives were decontextualized or split into standalone story modules. Participants picked the story modules from a given story library to form up a storyline. Although the story modules in story libraries were decontextualized, they became contextualized again once they were inserted into virtual environments as a coherent storyline. As a result, training objectives were contextualized according to the context set by trainees, by which tailored training was delivered.

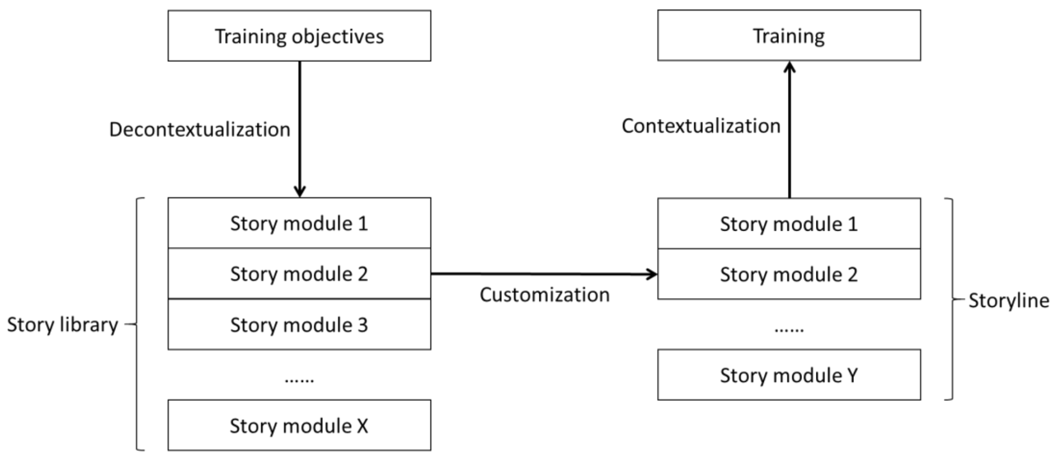


Figure 5-13 Decontextualization and contextualization of training objectives

In terms of pedagogy, the main purpose of this paper is not to discuss the effectiveness of the configurations of pedagogy applied to different participants, which will be presented in further papers. However, we still shed some light on the usability of teaching methods applied in the present study. The scores of teaching methods (Children: M = 3.53, SD = 0.563; Adults: M = 2.26, SD = 0.795) show that the pedagogy applied was well accepted by participants. Teaching methods were introduced as the last layer of information on top of context and representation. The main tone of this layer was to educate trainees through interactive ways, which involved visual content such as text-based instructions and feedback, and sound effects enhancing stimulation. The intervention of teaching methods did not alter storylines and virtual environments. It was a standalone dimension which could be tuned to utilize the utmost of each teaching method.

An overarching theme measured by our pilot trials was the engagement of the training. According to the results relating to engagement (Children: M = 3.53, SD = 0.524; Adults: M = 2.35, SD = 0.667), participants acknowledged the training itself was appealing, and they could easily focus on it. This finding is in line with that of previous studies where

IVR SGs were ad-hoc scenarios without customization (Feng et al., 2018). This is not an unexpected outcome as our customization process required little effort from participants, and thanks to the layered structure, the basic principles of IVR SGs were untouched. The outcome of customization (i.e., the training program) still held the essential principles, which were immersion, interaction, and user involvement, that assured an engaging IVR experience (Freina & Ott, 2015).

5.6. Conclusions

Adaptive game-based learning has been discussed over the decades. However, little attention has been paid to IVR-based earthquake emergency training, especially for the customization domain. In order to fill this gap, we proposed a user-based customization framework, which covers four dimensions of game-based learning and six parameters for tuning. Learner specifics, context, representation, and pedagogy are the four fundamental dimensions. The first research question was answered in this paper by identifying six parameters to effectively customize the IVR SG, including trainee type, virtual environments, earthquakes and damage, storylines, storytelling methods, and teaching methods. In response to the second research question, these six parameters contribute to the customization of four dimensions of an IVR SG according to the concept of game-based learning (see Figure 5-1 and Figure 5-5). Regarding the third research question, these parameters are suitable and competent to deliver a customized earthquake training experience for optimum learning. A case study was conducted, in which the framework was implemented, and a prototype was tested with end-users. Positive experimental results indicate that the customization process was well accepted by trainees, and the customization framework was competent to deliver a customized training program, with each fundamental dimension functioning. In future research, it is valuable to test the IVR SG training program against a non-IVR SG training program that is on flat screens controlled by keyboard and mouse.

IVR SGs have been applied to emergency and evacuation studies in recent years (Lin et al., 2019; Lin et al., 2020; Lovreglio et al., 2018). This paper dealt with the complex evacuation process related to earthquake emergencies, amalgamating different domains such as behavioral responses to emergencies, physical and perceptual complexity that makes up an indoor earthquake emergency, pedagogy and education aspects, and the computational features of IVR and SG into one unified training framework with customization features. In particular, this research contributes to knowledge in the domain of IVR SG applied to emergency evacuation by incorporating game-based learning into the customization framework. As such, this provides promising directions and opportunities to future research when it comes to the development of highly effective IVR SG-based training and education frameworks to enhance evacuation preparedness in the event of not only earthquakes, but also other natural and man-made hazards.

This chapter proposes the customization framework (see Figure 5-1) and validates the usability of a customizable IVR SG through case studies. The pedagogical impacts of the teaching methods and narrative methods in the customization framework are discussed in Chapter 6 and 7.

Chapter 6: Teaching Methods in the Customization Framework

The content of this chapter is extracted from:

Feng, Z., González, V.A., Mutch, C., Amor, R., & Cabrera-Guerrero, G. (2020). An Immersive Virtual Reality Serious Game to Train Children in Earthquake Emergency Responses. (Submitted to *Journal of Computer Assisted Learning*)

A customization framework (see Figure 5-1) for IVR SGs suited to earthquake emergency training has been developed in Chapter 5. In this chapter, the pedagogical impacts of the teaching methods in the customization framework are investigated. The participants of the experimental group in this chapter were the same as those in Chapter 5 (students).

6.1. Introduction

Earthquakes and the associated post-earthquake evacuation are dangerous and complex, and children are very vulnerable in these extreme situations (Cahill, Beadle, Mitch, Coffey, & Crofts, 2010; Lori, 2008; Mutch & Marlowe, 2013). It is estimated that about 25,000 children were killed worldwide over the decade between the late 20th century and early 21st century due to school collapse in earthquakes (Kenny, 2009). Schools have been found to be ill-prepared to protect children from such disasters (Mutch, 2014; Sapien & Allen, 2001; Tatebe & Mutch, 2015). Children, however, can play a fundamental role in enhancing disaster mitigation and risk reduction in communities by helping educate their families and friends (Cahill et al., 2010; Izadkhah & Hosseini, 2007). In order to build resilient families and communities, it is important, therefore, to increase children's disaster awareness, understanding and preparedness (Izadkhah & Hosseini, 2005; Mutch, 2014; Mutch, 2018). Children can be equipped with appropriate earthquake safety and emergency evacuation procedures through education and training in schools (V. A. Johnson, Ronan, Johnston, & Peace, 2014; Mutch, 2014).

Traditionally, children learn earthquake safety knowledge through textbooks, classroom lessons, and safety drills (Izadkhah & Hosseini, 2007). However, these methods have been criticized as poorly implemented and ineffective in enhancing children's disaster preparedness (Victoria A. Johnson, Johnston, Ronan, & Peace, 2014; Ramirez, Kubicek, Peek-Asa, & Wong, 2009; Ronan, Alisic, Towers, Johnson, & Johnston, 2015). Suggestions for preparedness enhancement include life-like exercises, less-familiar-scenarios practice, and the integration of theoretical and practical activities (Codreanu, Celenza, & Jacobs, 2014; V. A. Johnson et al., 2014; Ramirez et al., 2009). Much attention has been paid to practice, but it is dangerous and unethical to put children in a training situation with risky scenarios and hazardous events.

One possible alternative is using Immersive Virtual Reality (IVR) Serious Games (SGs) (Feng et al., 2020). IVR is a digital technology that allows users to be immersed in a simulated environment (LaValle, 2016). The credible experience makes users feel they are physically in the simulated environment (Sherman & Craig, 2018). SGs are video games whose main purpose is not entertainment, but training, education, or healthcare (Michael & Chen, 2006; Susi et al., 2007). IVR SGs facilitate practice in extreme and dangerous situations, without worrying about causing harm or generating large ethical issues. Note that ethical issues may still remain as trainees can get stressed and frightened in realistic virtual environments; nevertheless, it is better than exposing trainees with real dangerous events.

Our paper explores using IVR SGs to train children about how to respond effectively to an earthquake emergency. Research questions include: 1. How can children be trained in earthquake emergency responses using IVR SGs? 2. What are the impacts of the IVR SGs targeting children? A problem-based gaming framework was applied as the guideline to develop the proposed IVR SG training system. Children were exposed to an indoor earthquake and post-earthquake evacuation situation. Through solving problems in the simulation, children were trained to gain safety knowledge and to enhance their self-efficacy to cope with an earthquake and post-earthquake evacuation. Three teaching approaches were deployed to facilitate learning: immediate feedback, post-game assessment, and prior instruction. The effectiveness of these approaches was tested against a traditional training method (control group), which was the reading of a leaflet. Based on the discussion of the results, further improvement of the IVR SG training system and future research directions were identified.

6.2. Background

This section presents a review of IVR SGs for children in the literature. Then, a problem-based gaming framework is discussed, guiding the design of the proposed IVR SG training system.

6.2.1. IVR SGs for children

IVR SG applications for children have become popular in recent years. Many of them focus on medicine or healthcare, including physical and psychological assessment, treatment and intervention (Didehbani, Allen, Kandalaft, Krawczyk, & Chapman, 2016; Ip et al., 2018). In addition, a few of them have paid attention to safety training and safety behavioral study, targeting children (Morrongiello, Corbett, Milanovic, Pyne, &

Vierich, 2015; Schwebel et al., 2016; S. Smith & Ericson, 2009). Schwebel et al. (2016) proposed a projection-based IVR system (IVR is fulfilled by projecting video simulation on surrounding screens) to train children in pedestrian safety. Children of seven and eight years old received street-crossing training through the system. Results indicated that children's decision-making about the safety of traffic gaps was improved after the training. Similarly, Morronegiello et al. (2015) assessed child pedestrian behavior with a head-mounted display IVR system. Results showed that children from seven to ten years old performed evasive actions if they perceived risky conditions when crossing streets. In terms of emergency training, S. Smith and Ericson (2009) put children from seven to eleven years old in a Cave Automatic Virtual Environment (CAVE)-based training system to teach fire-safety skills. Results showed that the IVR training system did not affect short-term learning gains. However, it was found that the IVR training system was beneficial to the improvement of motivation for fire-safety training, as children were more engaged by having fun during the training. This finding is consistent with that of Virvou and Katsionis (2008) who argued that SGs were highly motivating for children to learn.

Previous studies support the notion that IVR SGs are promising to support education and training for children. While street-crossing and fire safety have been studied, there has been little discussion about earthquake-safety training. This study proposes an IVR SG training system, which targets children to learn and practice the behavioral responses to earthquakes and post-earthquake evacuation.

6.2.2. Problem-based gaming

The proposed IVR SG training system (see Section 6.3) was developed based on a problem-based gaming (PBG) framework (Kiili, 2007). The principles of PBG originated from problem-based learning (PBL), by which learners acquire knowledge by solving authentic problems through practice (Walker & Shelton, 2008). In the context of video games, authentic problems are contextualized as training tasks through game environments and storylines (Kiili, 2007). Trainees acquire knowledge and skills through the experience of accomplishing these tasks. Figure 6-1 shows the concept of the PBG model.

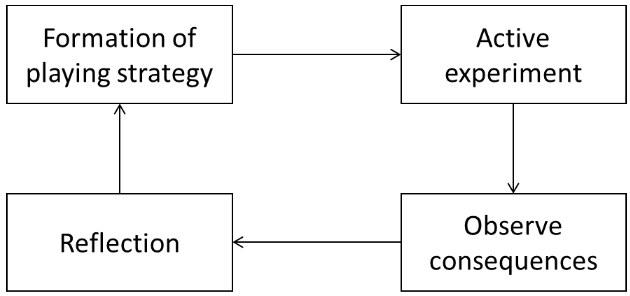


Figure 6-1 Problem-based gaming, derived from (Kiili, 2007)

As discussed by Kiili (2007), the first phase of PBG starts with strategy formation, by which trainees make up a strategy to solve follow-on problems. Usually, the strategy is built up based on trainees' prior knowledge and experience. Then, trainees apply the strategy and observe the consequences of their actions. Next, the reflection phase takes place, in which trainees revisit their previous action and think over it (Boud, Keogh, & Walker, 2013). Reflection is a critical phase of the PBG as the outcome of reflection usually lies in the construction and synthesis of knowledge (Kiili, 2007), which is the ultimate goal of an IVR SG training system.

Reflection can be facilitated in different ways. A commonly applied approach is to provide feedback to trainees (Kiili, 2007). Feedback is given to trainees to inspire reflective thinking and construct knowledge (C. I. Johnson, Bailey, & van Buskirk, 2017). With the integration of IVR SGs, two common types of feedback have been applied widely, namely immediate feedback and post-game assessment (i.e., delayed feedback)

(Feng et al., 2018). The main differences between these two forms of feedback lie in the sets of information provided and associated timing. Immediate feedback means the feedback is provided to trainees immediately after they undertake a training task as defined by the storyline. Chittaro and Buttussi (2015) deployed text-based and audio-based immediate feedback to trainees when they took a wrong action in an aviation safety training program. Mistakes were warned by immediate feedback. Reflection could occur when trainees received such information that ultimately reflects what went wrong. Consequently, appropriate safety knowledge was constructed. Statistically significant knowledge gain was observed Chittaro and Buttussi (2015). Regarding post-game assessment, it is given to trainees at the end of an entire training storyline. Thus, trainees can go through the IVR SG storyline and make their responses to training tasks in a flow state, without interruption. Li et al. (2017) provided trainees with an assessment after an earthquake emergency, which was obtained based on trainees' overall performance during the earthquake, regarding how well they protected them against hazards. Hence, trainees were encouraged to reflect and evaluate their previous behavioral responses after the completion of the training storyline. As a result, self-protection skills were developed with the assessment of previous performance. Li et al. (2017) reported that the participants trained by their IVR SG had developed significantly better self-protection skills than those trained by videos or manuals.

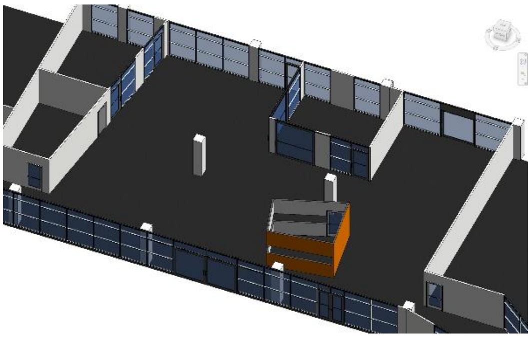
In the context of IVR SGs, both forms of feedback have been argued to be effective to enhance trainees' preparedness for emergency situations. However, these studies were targeting adults, and little attention has been paid to children. Children have a different learning process in comparison with adults (Kuhn & Pease, 2006). In a vocabulary learning study, delayed feedback was found to be more effective for children (Grade 6 students) than immediate feedback (Metcalfe et al., 2009). One possible explanation was that with immediate feedback, incorrect responses were still fresh in children's mind, which could interfere with the acquisition of correct responses (Kulhavy & Anderson, 1972). As for delayed feedback, children were more likely to forget their incorrect responses so that less interference would occur. Another possible reason could be the double exposure brought by delayed feedback (Clariana, Wagner, & Murphy, 2000). The first exposure occurs when children are facing a problem, and the second exposure happens when the delayed feedback allows children to revisit the previous problem after a period (Clariana et al., 2000). However, there is a debate about the effectiveness of immediate and delayed feedback for children. Studies report that immediate feedback is more effective when teaching in classrooms (Kulik & Kulik, 1988). A meta-analysis also suggests that immediate feedback is more beneficial to lower-order learning (i.e., recognize and understand concepts) than delayed feedback (van der Kleij, Feskens, & Eggen, 2015). The controversial situation leads to an interesting topic of how different forms of feedback perform in the case of training children in earthquake emergency responses by IVR SGs. Since little discussion has been made on this topic, we decided to test both forms of feedback in our proposed IVR SG training system.

As well as feedback, another alternative to encourage reflection is prior instruction. Prior instruction helps trainees focus on relevant information and establish inferential links between their prior knowledge and learning content (McCrudden, Schraw, & Kambe, 2005; McCrudden, Schraw, & Hartley, 2006; McCrudden, Magliano, & Schraw, 2010). Erhel and Jamet (2013) argue that prior instruction plays an important role in facilitating learning motivation in the context of digital game-based learning. Similarly, in an IVR SG fire safety study which targeted children from 7 to 11 years old, S. Smith and Ericson (2009) note that prior instruction helps improve trainees' motivation. Reflection occurs when trainees are retrieving prior instruction after dealing with learning tasks, and is promoted by a higher level of learning motivation (R. E. Mayer & Johnson, 2010). Motivation is the main driving force to facilitate reflection, which also happens to be one of the features enabled by IVR SGs (Huang, Rauch, & Liaw, 2010; McCrudden et al., 2010). Hence, we decided also to test prior instruction with children in the case of IVR SGs-based earthquake emergency training.

6.3. The IVR SG training system

The IVR SG training system simulates an earthquake occurring at Ormiston Junior College, a junior secondary school located in Auckland, New Zealand, from a first-person perspective. Students from the school were recruited as participants to test the IVR SG training system.

Plan drawings and on-site 360-degree panoramas were referred to as the data source to develop the virtual environment. The virtual environment was developed following a Building Information Modelling (BIM)-based workflow, by which accurate building layout can be defined, and dynamic changes can be applied to individual objects to enable credible earthquake simulations (Feng et al., 2018; Lovreglio et al., 2018). A basic BIM model was developed using Autodesk Revit (see Figure 6-2), including walls, ceilings, floors, columns, doors and windows. This model determined the envelope and layout of the built environment, where the training would take place. Then the model was imported into Unity for further development. Low-polygon models (mesh models with small numbers of polygons) of furniture and appliances were placed in the building model, and IVR and game mechanisms were developed at this stage.



(a) The model in Revit, without furniture and appliances (ceilings are hidden for demonstration).

(b) The final model in Unity, including furniture and appliances (ceilings are hidden for demonstration).

Figure 6-2 School models developed in Revit and Unity.

The IVR SG training system aims to provide a training simulation instead of an actual physical performance modelling and analysis of the building stressed by an earthquake. In this case, the earthquake and its damage were modelled following a qualitative approach, by which the simulation was developed based on qualitative descriptions (Lovreglio et al., 2018). The qualitative approach gives trainees a perception of an earthquake with emotional arousal through a credible experience. Geonet (New Zealand's national geophysical monitoring system) provides a database with detailed qualitative descriptions based on the New Zealand Modified Mercalli Intensity (MMI) Scale (GeoNet, 2019). According to the database, we selected MMI 6 as the reference to build the earthquake simulation, in which case "Furniture and appliances may move on smooth surfaces, and objects fall from walls and shelves. Glassware and crockery break. Slight non-structural damage to buildings may occur" (GeoNet, 2019). The reason to select MMI 6 is that it represents a strong earthquake where buildings do not collapse. Given this, it is feasible to set up a training program (e.g., post-earthquake evacuation training) in a damaged building, and trainees can get a decent experience of earthquakes. Along with the visual simulation of earthquakes, sound effects were also deployed in order to increase the credibility of the simulation, including rattling sounds, creaking sounds, breaking sounds, and human sounds (e.g. screams, whining).

During an earthquake and post-earthquake evacuation, people need appropriate behavioral responses in order to minimize injuries or casualties as recommended by best practice. New Zealand Civil Defence has issued an emergency guideline called "What's the Plan Stan?", specifically targeting children (Ministry of Civil Defence and Emergency Management, 2019). This guideline recommends a list of behavioral responses to an earthquake and post-earthquake evacuation. Based on that, we defined seven training objectives, which are the desired knowledge to be imbued into trainees after training, as shown in Table 6-1.

Table 6-1 Recommended behavioral responses as training objectives

Phase	Behavioral responses
During an earthquake	Drop, Cover and Hold
	Stay where you are until the shaking stops
	Do not attempt to run outside
After an earthquake	Listen to and follow all instructions from adults or the radio
	Stay calm. If you can, help others who may need it
	Watch out for possible hazards
	Remember there may be some aftershocks

The storyline of the IVR SG training system was established based on the defined training objectives. A storyline consists of a series of scenarios where knowledge is conveyed to trainees (Starks, 2014). According to PBG, problems were developed in each scenario using the training objectives as a reference (see Table 6-2). With exposure to the scenarios, trainees were expected to solve the problems and learn from them.

Table 6-2 The scenarios and problems

Phase	Scenarios	Problems to solve
Before an earthquake	Trainees start walking in a corridor, which leads to a large common area.	No
	Once trainees reach a desk in the common area, an earthquake strikes the building where trainees stay.	No
During an earthquake	Trainees can choose to take cover under a table, or beside a shelf, or run out of the building.	Yes
After an earthquake	A teacher approaches trainees to ask them to follow her instructions. Trainees can choose to follow her or run out of the building.	Yes
	An aftershock hits the building. Trainees can choose to take cover under a table, or beside a cabinet, or run out of the building.	Yes
	Once the aftershock stops, trainees start to leave the building.	No
	On the way out, trainees need to detect hazards and decide whether to avoid them.	Yes
	A student is trapped under a table. Trainees can choose whether to offer help.	Yes

The storyline was narrated on an action-driven basis, where trainees needed to take actions to solve the problems so that they could progress through the storyline (Feng et al., 2018). Once trainees had taken action, the storyline progressed linearly to the next scenarios, with no possibility to revisit a previous scenario or revise a previous response to a problem.

As discussed in Section 6.2.2, three different approaches to promote reflection were applied to our IVR SG training system involving prior instruction, immediate feedback, and post-game assessment. Each approach represents a unique way to encourage reflection for trainees, as shown in Figure 6-3.

(a1) Prior instruction was given to trainees before a problem

(a2) After receiving instructions, trainees experienced an event and tried to select an action choice in response to the event

(b1) Trainees experienced an event and tried to select an action choice in response to the event

(b2) Immediate feedback was given to trainees after solving the problem

(c1) After the entire storyline, a checklist is presented to trainees as post-game assessment

(c2) Trainees revisited their previous training experience through the post-game assessment

Figure 6-3 The teaching approaches of the IVR SG training system: (a) Prior instruction, (b) Immediate feedback, (c) Post-game assessment

Prior instruction was given to trainees before the problems they were about to face in a text-based format. The content of the instructions included the recommended behavioral responses to the problems. Trainees read the instructions first before solving the problems. For the purpose of validating our research hypothesis (see Section 6.4.1), no further information was presented to trainees once they had tried to solve the problems.

Regarding immediate feedback, a panel with a green check or a panel with a red cross was presented to trainees according to the responses they had made to solve the problems, indicating whether their responses were correct or not. Simultaneously, two different sound effects were triggered respectively in order to enhance the stimulation of feedback, with “Ding” for a green check and “Bang” for a red cross. Trainees only received feedback immediately after they solved a problem.

In terms of post-game assessment, once the entire storyline was over, trainees were led through a playback that reported their responses made in previous training against the recommended behavioral responses listed in Table 6-1. Trainees only reviewed their responses at this stage, and no other information was presented before and after trying to solve the problems during the training.

6.4. Research Methods

There are three versions of the proposed IVR SG training system, with three teaching approaches respectively. We undertook an experimental study to assess the three IVR SGs against a traditional method (reading a leaflet). The study was guided by a pretest-posttest research design. A pretest measure on the interested outcomes was conducted prior to intervention, with the same measure administered after training (Salkind, 2010). This experimental design is good at exposing the effects of a treatment or an intervention on a group of subjects. We conducted within-groups analysis to investigate the effects of each training approaches. We also applied between-groups analysis to compare these training approaches regarding the effectiveness of knowledge gain and self-efficacy improvement. In the following section, the hypothesis of the study is outlined first, followed by the information of materials and participants (i.e., trainees). Next, the measures administered before and after the training are discussed. Lastly, the procedure to carry out the study is described.

6.4.1. Hypotheses

As discussed in Section 6.2, previous IVR SGs studies with three different teaching approaches have shown promising outcomes to enhance the emergency preparedness of trainees. However, the effectiveness of these approaches for children in IVR SGs-based earthquake emergency training remains unclear. Therefore, we proposed the following hypothesis in the present research:

An IVR SG training system with any teaching approach (immediate feedback, post-

game assessment, and prior instruction) would be more effective than a traditional training material (i.e., a leaflet) to train children on appropriate earthquake emergency responses.

The effectiveness of a training method lies in the capability to improve the preparedness of trainees, such as knowledge acquisition, knowledge retention, self-efficacy improvement, attitude changes, and skills transfer (Tekeli-Yeşil et al., 2010). This study measures two dimensions: knowledge acquisition and self-efficacy improvement. In this study, it was hypothesized that an IVR SG training system would be more effective in terms of knowledge acquisition and self-efficacy improvement.

6.4.2. Equipment

The IVR SG training system was implemented with Unity and run on a DELL PC workstation with an Intel® Xeon® W-2125 processor, an NVidia GeForce RTX 2080 graphics card, and 64 GB ram. The IVR was delivered using an Oculus Rift VR system, which includes a head-mounted display (HMD) with 1080x1200 resolution per eye and a 110-degree field of view, two remote controllers, and two tracking sensors. The graphic output of the HMD was streamed on an LED monitor which was connected to the same PC workstation, allowing observation in real-time of the trainees' responses in the IVR SG training system.

6.4.3. Trainees

One hundred and twenty-five secondary school students (46 girls and 79 boys) with ages ranging from 11 to 15 years old ($M = 12.4$, $SD = 1.08$) from Ormiston Junior College (a junior secondary school in Auckland) took part in the study. Students were recruited by their teachers and from referrals from their friends. The study was approved by The

University of Auckland Human Participants Ethics Committee, and the board of trustees of Ormiston Junior College.

Trainees were randomly assigned to four groups. Three groups were trained with the IVR SG training system (immediate feedback, 31 trainees; prior instruction, 32 trainees; post-game assessment, 31 trainees) while the last group was a control group trained through studying a leaflet (31 trainees).

Trainees' previous experience involving earthquake drills and IVR was collected, as shown in Table 6-3. Kruskal-Wallis tests were conducted confirming no statistically significant differences among the four groups (earthquake drills, $p = 0.526$; IVR, $p = 0.276$).

Table 6-3 The numbers of trainees who had previous experience with earthquake drills and IVR

Groups (total trainees)	Earthquake drills		IVR	
Immediate feedback (31)	17	54.8%	18	58.07%
Prior instruction (32)	22	68.75%	24	75.00%
Post-game assessment (31)	16	51.61%	19	61.29%
Leaflet (31)	19	61.29%	16	51.61%

Since the participant are the same as those in Chapter 5 (students), the adequacy of sample size and power analysis has been discussed in Section 5.4.2.

6.4.4. Measures

In order to validate the hypothesis, two factors were measured against the outcome of training: safety knowledge and self-efficacy.

6.4.4.1. Safety knowledge

The earthquake safety knowledge learned through training should be applicable to tackle similar real-life scenarios (Li et al., 2017). In our case, the safety knowledge about appropriate behavioral responses is essential to guide evacuation behaviors in an earthquake emergency. We assessed the acquisition of the safety knowledge regarding the appropriate behavioral responses taught by the IVR SG training system. A knowledge test containing eight questions was undertaken before and after the training. The knowledge test was in the format of true-false statements. The reason to use this format is that specific questions are more likely to prompt the accuracy of the information reported by children than open-ended questions (Peterson, Dowden, & Tobin, 1999). Trainees were asked to pick out true statements only. Possible scores ranged from 0 to 8, where trainees lost one mark if they failed to identify a true statement or selected a false statement which they thought it was true. Table 6-4 shows the statements and the correct answers to the knowledge test.

Table 6-4 Knowledge test

Statements	True or false
You can start running outside when an earthquake is hitting the building you are in.	False
You need to follow instructions from teachers.	True
You can take cover under a table during an earthquake.	True
You can start evacuation if you feel unsafe during the shaking.	False
Aftershocks can happen after a major shock.	True
You can keep running outside if an aftershock is hitting.	False

If hazardous objects such as debris and broken glass are in your way, you can ignore them to speed up evacuation.	False
Do not care about others because your personal safety comes first.	False

6.4.4.2. Self-efficacy

Self-efficacy is people's beliefs in their ability to accomplish tasks, overcome challenges, or influence events (Bandura, 1982; Bandura, 2010). Self-efficacy is an indicator to shed light on behavior change and performance improvement (Bandura & Adams, 1977; Bandura, 1977; Stajkovic & Luthans, 1998). Self-efficacy can be measured by the General Self-Efficacy Scale, which includes several self-report statements (Schwarzer & Jerusalem, 2010). Based on this general scale, we designed six statements for trainees to answer before and after the training:

1. "I know what to do when facing an earthquake";
2. "I can remain calm when facing an earthquake";
3. "I have the confidence to deal with an earthquake emergency";
4. "I can come up with a plan for responses to an earthquake";
5. "I can handle situations during an earthquake";
6. "I can think of a solution if I am in trouble during an earthquake."

According to the General Self-Efficacy Scale, trainees were asked to choose their levels of agreement to each statement from "Not true", "Hardly ever true", "Sometimes true", and "Always true". This response format is suitable for children, as studies have suggested that children prefer words to numbers to reflect and describe their thoughts (Mellor & Moore, 2013). Answers were transcribed to scores ranged from 1 to 4 where 1 stood for "Not true", and 4 stood for "Always true". The final score of a trainee was obtained by getting the sum of all statements, in which case final scores ranged from

6 to 24 (Schwarzer & Jerusalem, 2010). Trainees answered the self-efficacy statements before and after the training.

6.4.5. Procedure

The present experiment was carried out in a standalone room at Ormiston Junior College. Trainees, their parents and teachers read the Participation Information Sheet and signed consent forms. Trainees could quit the experiment at any time without any reason. Upon arrival, trainees were randomly assigned to four groups (see Table 6-3). Next, trainees answered a questionnaire which included demographical information, a knowledge test, and self-efficacy statements.

After the completion of their questionnaire, trainees in IVR groups were asked to put on an IVR headset. Trainees were assisted in getting a clear view with the IVR headset. Trainees were led through a tutorial which helped them get familiar with controls and navigation in IVR environments. Once they were comfortable with it, the IVR training session took place. For the IVR groups, the only difference between groups was the teaching approach being used.

For the control group, after the pre-test questionnaire, trainees were treated with one of the traditional training methods, in which case, was self-learning through leaflets. Safety knowledge was presented on leaflets in the form of texts. This leaflet was a printed version of a webpage which was originally designed by New Zealand Civil Defence, called “What’s the plan, Stan?” (Ministry of Civil Defence and Emergency Management, 2019). “What’ the plan, Stan” is a webpage targeting children and encourages children to learn safety knowledge for various types of emergencies. Trainees were instructed to study the leaflet carefully by themselves. No other types of instructions and explanations were provided. Trainees could spend as much time as they wanted until they fully understood the content in the leaflet.

Lastly, when trainees completed training sessions, they answered a questionnaire which included a knowledge test and self-efficacy statements.

6.5. Results

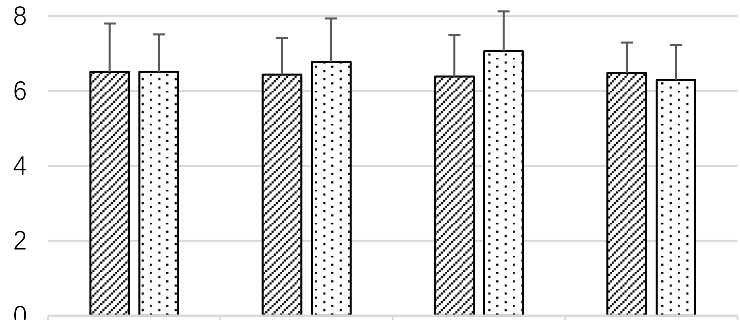
Figure 6-4 shows the comparisons of the mean values of knowledge scores and self-efficacy scores. For the immediate feedback group, the mean pre-test knowledge score was 6.52 (SD = 1.29), and the mean post-test knowledge score was 6.52 (SD = 1.00), with no statistically significant improvement found; the mean pre-test self-efficacy score was 16.39 (SD = 3.853), and the mean post-test self-efficacy score was 17.94 (SD = 3.898), with statistically significant improvement found ($p = 0.001$, Wilcoxon Signed Ranks Test).

For the post-game assessment group, the mean pre-test knowledge score was 6.39 (SD = 1.12), and the mean post-test knowledge score was 7.06 (SD = 1.06), with statistically significant improvement found ($p = 0.018$, Wilcoxon Signed Ranks Test); the mean pre-test self-efficacy score was 15.97 (SD = 4.301), and the mean post-test self-efficacy score was 18.65 (SD = 3.971), with statistically significant improvement found ($p < 0.001$, Wilcoxon Signed Ranks Test).

For the prior instruction group, the mean pre-test knowledge score was 6.44 (SD = 0.98), and the mean post-test knowledge score was 6.78 (SD = 1.16), with statistically significant improvement found ($p = 0.029$, Wilcoxon Signed Ranks Test); the mean pre-test self-efficacy score was 15.66 (SD = 3.633), and the mean post-test self-efficacy score was 17.38 (SD = 3.933), with statistically significant improvement found (p = 0.002, Wilcoxon Signed Ranks Test).

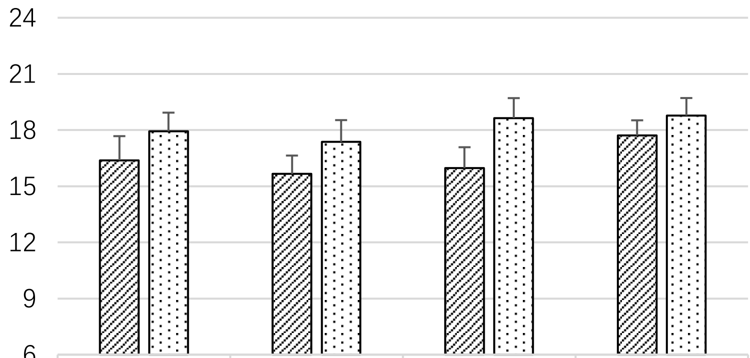
For the leaflet group, the mean pre-test knowledge score was 6.48 (SD = 0.81), and the mean post-test knowledge score was 6.29 (SD = 0.94), with no statistically significant improvement found; the mean pre-test self-efficacy score was 17.71 (SD = 3.977), and the mean post-test self-efficacy score was 18.77 (SD = 3.283), with statistically significant improvement found (p = 0.047, Wilcoxon Signed Ranks Test).

Safety knowledge



	Immediate Feedback	Prior Instructions	Post-game Assessment	Leaflet
Pre-test	6.52	6.44	6.39	6.48
Post-test	6.52	6.78	7.06	6.29

Self-efficacy



	Immediate Feedback	Prior Instructions	Post-game Assessment	Leaflet
Pre-test	16.39	15.66	15.97	17.71
Post-test	17.94	17.38	18.65	18.77

Figure 6-4 Mean pre-test and post-test knowledge and self-efficacy scores. Error bars indicate the standard deviation of the mean

In order to deal with the pre-test and post-test scores between four treatment groups, we analyzed knowledge scores and self-efficacy scores using ANCOVA (B. H. Cohen, 2013). Pre-test scores were served as the covariate, and post-test scores were served as the dependent variable. The analysis controls for pre-test scores. Results indicated significant difference between groups in the post-test scores of knowledge ($F(3,120) = 4.218$, $p = 0.007$, $\eta_p^2 = 0.095$), and no significant difference between groups in the post-test scores of self-efficacy ($F(3,120) = 1.208$, $p = 0.310$, $\eta_p^2 = 0.029$). Follow-on pairwise comparisons were carried out using Bonferroni tests. Comparison results are reported in Table 6-5.

Table 6-5 Between-group comparisons

Pairwise comparisons	Knowledge acquisition	Self-efficacy improvement
Leaflet vs. Immediate feedback	No significant difference	No significant difference
Leaflet vs. Prior instruction	No significant difference	No significant difference
Leaflet vs. Post-game assessment	Significant difference ($p = 0.007$)	No significant difference
Immediate feedback vs. Prior instruction	No significant difference	No significant difference
Immediate feedback vs. Post-game assessment	No significant difference	No significant difference
Prior instruction vs. Post-	No significant difference	No significant difference

game assessment

6.6. Discussion

The results of the experiment partially support our hypothesis (see Section 6.4.1). Considering the knowledge acquisition between pre-test and post-test, the IVR SG training system with post-game assessment showed better performance than the leaflet approach. However, no performance differences were reported between the leaflet and the IVR SG training system with immediate feedback or prior instruction. Regarding self-efficacy improvement between pre-test and post-test, the leaflet and IVR SG training system with post-game assessment, immediate feedback, or prior instruction were all effective and produced similar results.

Based on ANCOVA and Bonferroni tests, no significant differences regarding knowledge acquisition and self-efficacy improvement were reported between the IVR SG training system with immediate feedback, post-game assessment, or prior instruction. However, we cautiously argue that their performance was not similar. According to within-groups analysis (Wilcoxon Signed Ranks tests), post-game assessment and prior instruction showed a significant effect on knowledge acquisition while immediate feedback was not significantly effective. The trend of this difference is shown in Figure 6-5. No significant between-groups difference does not mean no difference (Gelman & Stern, 2006). In our experiment, these three approaches are not significantly different, but we are uncertain how this will play out in the general population (children).

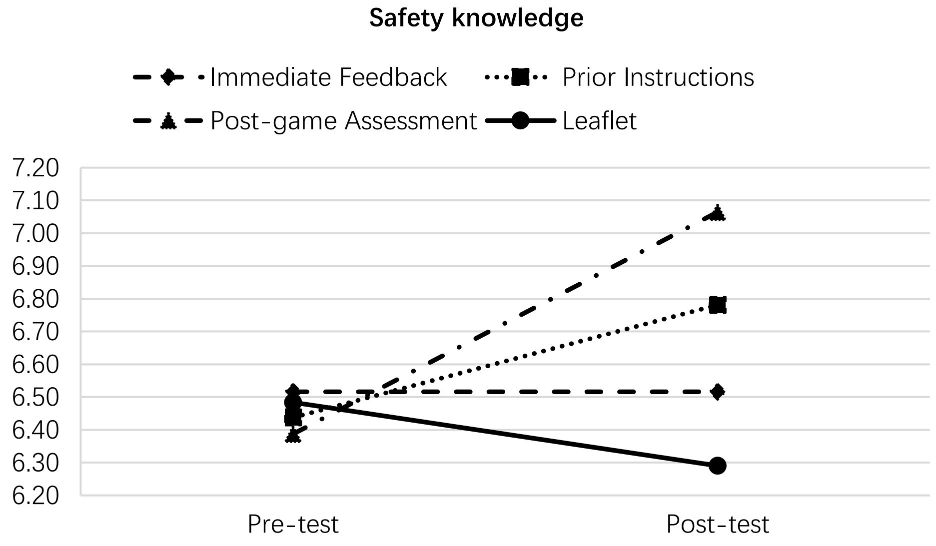


Figure 6-5 The trend of pre-test and post-test knowledge scores

As compared to the leaflet approach, the IVR SG training system with post-game assessment is the most effective tool among the three IVR conditions to train children in earthquake emergency responses. Within-group positive knowledge gain and self-efficacy improvement were observed after the training. Trainees received a checklist and playback regarding their performance in the simulated earthquake and post-earthquake evacuation, in which case trainees have been led through the whole response procedure again. Trainees were expected to reflect on their previous behavior against the feedback they were taking. The second exposure brought by post-game assessment could result in enhanced memory for knowledge gain (Clariana et al., 2000). Our results, consistent with the findings from other earthquake safety training in IVR SGs (Feng et al., 2020; Li et al., 2017), supports the possibility that IVR SGs with post-game assessment can help children get prepared for earthquakes and post-earthquake evacuation.

We did not obtain the expected results with respect to immediate feedback in comparisons with leaflets. Within-group positive self-efficacy improvement was discovered, while no knowledge gain was found. The proposed IVR SG training system was designed to put trainees in an earthquake and follow-on post-earthquake evacuation, keeping them in a state of flow (Admiraal et al., 2011). Trainees went through scenes from one to another without interruption, as in an actual earthquake emergency. However, this seemed to bring a problem. Trainees had little time to reflect as the experience flow was not paused when they were receiving immediate feedback. The insufficient reflection time could result in inadequate reflection, which could lead to poor performance for knowledge gain (Romano & Brna, 2001). Trainees might get distracted from reflection when visual simulations and sound effects were still ongoing, especially for children, who are more likely to lose attention than adults (Matusz et al., 2015). Our results differ from the findings of other studies (Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015). Burigat and Chittaro (2016) applied an IVR SG with immediate feedback to train participants in the spatial knowledge of an aircraft to perform effective evacuation. No pauses (or reflection time) were available for participants after receiving feedback. A positive training outcome was reported in that study, with zero reflection time. Chittaro and Buttusi (2015) used an IVR SG with immediate feedback to teach participants about the behavioral responses of aircraft emergencies. A 7-second pause was given to participants after receiving feedback. A positive outcome was observed in this study. Taken together, it is possible to infer that the effectiveness of immediate feedback with or without dedicated reflection time relies on the type of knowledge to be conveyed. Research questions to be asked could include how reflection time after immediate feedback would impact on the effectiveness of teaching different types of knowledge through an IVR SG, such as factual knowledge or procedural knowledge (Krathwohl & Anderson, 2009).

Another possible explanation for the poor performance of immediate feedback is that we deployed the same level of stimulation mechanism for both positive and negative feedback. Positive feedback means the visual and acoustic stimulation given to trainees when they made correction actions, as with negative feedback for incorrect actions. The exposure to negative feedback may cause a negative suggestion effect, by which trainees remember misinformation as being correct after the training (A. S. Brown, Schilling, & Hockensmith, 1999). In our case, the possible explanations of this effect include memory impairment (i.e., incorrect information may harm the memory trace of the related correct information) and retrieval interference (i.e., the coexistence of correct and incorrect information in memory may make cause difficulties to retrieve the correct information) (A. S. Brown et al., 1999). It can thus be suggested that the stimulation mechanism for positive and negative feedback should be differentiated at different levels, with intense stimulation for positive feedback and light (or minimum) stimulation for negative feedback. This is an interesting issue for future research to see how the differentiation of feedback stimulation will influence the training outcomes of IVR SGs.

Another unanticipated finding was that the IVR SG training system with prior instruction did not show significantly better performance than the leaflet approach. This result was similar to the finding of a previous study, which used prior instruction to train children about fire safety knowledge (S. Smith & Ericson, 2009). No significant difference was revealed in that study comparing IVR and non-IVR groups. Even so, within-group positive knowledge gain and self-efficacy improvement were still produced in our study, which implies that an IVR SG with prior instruction is effective to enhance children's preparedness against earthquake emergencies. Trainees received prior instruction to guide their actions for the challenges of the training storyline. Trainees were expected to put recommendations into practice and learn from it. Once trainees had taken action in a scene, the IVR SG training system continued to progress to the next scenes. Unlike the immediate feedback condition, during the transition of scenes, trainees were not reading other information (i.e., immediate feedback), which gave them more time and room to reflect on what they had just done before. Our results support the notion that instructions have a positive influence on learning (McCrudden et al., 2005; McCrudden et al., 2006; McCrudden et al., 2010), even applicable with IVR SGs.

Surprisingly, a slight drop (not significant) of knowledge gain was found in the leaflet group. This may be due to the fact that information on the leaflet was simple guidelines, such as “Stay where you are until the shaking stop”, where a further explanation for this guideline was not given. Children appreciate explanatory information to help with learning rather than corrective information (Moreno, 2004). With the lack of explanatory information, children may develop a wrong perception of the content. In the proposed IVR SG training system, similar guidelines without explicit explanations were applied as prior instruction and feedback. However, through the experience of relevant scenes, actions, and corresponding consequences in the virtual environment, trainees perceived the explanatory information of guidelines in a different form. The credible simulation allowed trainees to practice events that were impossible to conduct in real life (e.g., hazardous events, life-threatening situations), as one of the advantages of IVR SGs (Sherman & Craig, 2018).

One interesting finding of our experiment was that the trainees of all groups improved self-efficacy significantly. No between-group difference was revealed. No matter whether trainees had practised in IVR SGs they perceived they were prepared for earthquakes and post-earthquake evacuation. However, there was a false judgement here. Only groups who were trained by the IVR SG training system with prior instruction and post-game assessment increased their safety knowledge related to suitable behavioral responses significantly. There was no significant knowledge gain for groups trained by leaflet and the IVR SG training system with immediate feedback. These two groups had developed false self-efficacy with incorrect knowledge bearing in mind. The results of our study do not explain the occurrence of this contrary situation, which leaves the room for future research to investigate.

6.7. Conclusions

The present study contributes to the current literature on extending the application of IVR SGs to children for earthquakes and post-earthquake evacuation training. The findings of previous IVR SGs studies for adults may not be applicable for children, as children process differently from adults in terms of learning (Kuhn & Pease, 2006). An IVR SG training system based on PBG was developed and explored, aiming to train children in earthquake emergency responses. Different teaching mechanisms were explored trying to find an effective approach suited for children, with a leaflet approach as a control group. Knowledge acquisition and self-efficacy improvement were measured. Results support the notion that IVR SGs with post-game assessment is more effective than the leaflet approach, which makes it the most effective to train children in IVR SGs. The performance difference was recognized between these three IVR SG groups. However, the difference was not significant, and we are uncertain how this will apply to the general population. Further research can keep investigating this point.

Although this study investigated different mechanisms for IVR SGs individually, this does not necessarily suggest that IVR SGs for children are only limited to one mechanism at a time. The integration of multiple mechanisms is worth investigating in the future to establish a robust model of IVR SGs to help children prepare for earthquakes and post-earthquake evacuation.

This chapter investigates the pedagogical impacts of different teaching methods proposed in the customization framework (see Figure 5-1). The investigation of narrative methods in the customization framework is discussed in the next chapter.

Chapter 7: Storytelling Methods in the Customization Framework

The content of this chapter is extracted from:

Feng, Z., González, V.A., Mutch, C., Amor, R., & Cabrera-Guerrero, G. (2020). Exploring Spiral Narratives with Immediate Feedback in Immersive Virtual Reality Serious Games for Earthquake Emergency Training. (Submitted to *Virtual Reality*)

A customization framework (see Figure 5-1) for IVR SGs suited to earthquake emergency training has been developed in Chapter 5. In this chapter, the pedagogical impacts of the narrative methods in the customization framework are investigated. The participants of the experimental group in this chapter were the same as those in Chapter 5 (adults).

7.1. Introduction

Earthquakes are one of the natural hazards that heavily impact people's lives, with approximately 100 significant earthquakes causing damage around the world every year (United States Geological Survey, 2018). Suitable and timely behavioral responses are essential to save lives and reduce injuries during earthquake emergencies (Alexander, 2012; Bernardini et al., 2016). Various training and educational approaches have been developed to disseminate safety knowledge about appropriate behavioral responses, including drills, seminars, and posters. In recent years, other alternatives, such as Immersive Virtual Reality Serious Games (IVR SGs), have increasingly gained attention for education and training (Freina & Ott, 2015). IVR offers credible virtual environments that engage users through a high level of immersion (LaValle, 2016; Sherman & Craig, 2018). In fact, users may develop a sense of presence that they are physically in the virtual world generated by IVR (LaValle, 2016; Sherman & Craig, 2018). IVR SGs offer the possibility to train people in life-like scenarios and simulations. Therefore, trainees have the chance to experience hazardous and extreme events, such as emergency situations, without being harmed or causing large ethical issues (Feng et al., 2018). Note that ethical issues may still remain, such as stressing and frightening trainees going through a highly engaging and credible IVR experience; however, it is more acceptable than making trainees face real danger. Risk-free training and exercise become available in close-to-reality simulations, with little compromise to ecological validity (Kinateder et al., 2014). In the literature, IVR SGs have been reported to deliver quality training for emergencies (Feng et al., 2018). Immersion, interaction, and engagement are the key features promoting learning outcomes (Freina & Ott, 2015; Li et al., 2017). As well as this, Krokos et al. (2019) also demonstrate that IVR improves memory recall by engaging trainees to focus on learning tasks.

IVR SGs are in line with the concept of problem-based gaming (PBG), where trainees develop skills and acquire knowledge by solving authentic problems through the learning and practice of using games (Kiili, 2007; Walker & Shelton, 2008). Figure 7-1 illustrates the concept of PBG. According to Kiili (2007), in the first phase, trainees start to form a strategy to solve the present problems based on their prior knowledge and experience. Next, trainees actively solve the problems and observe the consequences of their responses. Then, trainees reflect on their strategy and performance (Boud et al., 2013). Reflection is a vital part of PBG as it constructs cognition and synthesizes knowledge (Kiili, 2007). The trainees' behaviors continuing the game are determined by reflection, where trainees decide whether to stick with the previously formed strategy or alter it based on input from reflection (Kiili, 2007).

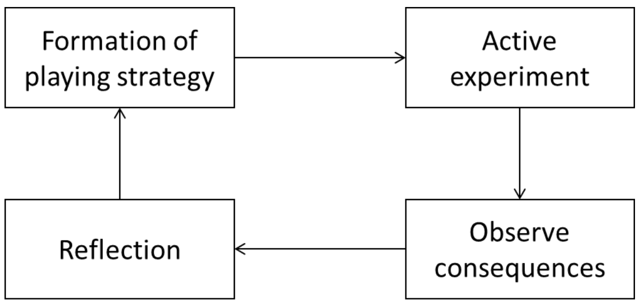


Figure 7-1 Problem-based gaming, derived from (Kiili, 2007)

PBG suggests that the interaction between trainees and games is an ongoing process causing a loop, with a focus on constructing knowledge and improving problem-solving strategies used by trainees continuously in the gameplay. With the current framework of PBG, after making a response, trainees observe consequences, engage in reflection, and carry on solving the next problems. However, the current PBG lacks the possibility for trainees to reflect while they are making a response. Trainees do not have a chance to adjust their responses and reshape their behaviors for the current scenario. In contrast, they only establish perspectives and knowledge with the reflection after the completion of the current scenario. As such, PBG has limited ability to enable reflection-in-action. Reflection-in-action refers to the reflection that takes place while individuals are carrying out an activity (Greenwood, 1993). Individuals need real-time feedback to make sense of confronting situations, adapt themselves, adjust behaviors, and take actions accordingly (Greenwood, 1993). Reflection-in-action shapes the output of individuals, in most cases, modifying their behavioral responses to current situations (Greenwood, 1993). Therefore, the facilitation of reflection-in-action is important to train individuals about behavioral responses to specific scenarios.

This study explores the possibility to integrate reflection-in-action and reflective redo within an IVR SG training system, targeting earthquakes and post-earthquake evacuation training. Research questions have been raised to guide the present study: 1. How can reflection-in-action and reflective redo be implemented in an IVR SG? 2. What are the impacts of such an IVR SG framework? A literature review on IVR SGs, storylines, and reflection is outlined first. Following that, a detailed discussion of the design and development of the proposed IVR SG training system is mapped out. Reflective redo and reflection-in-action are realized through two game mechanisms: a spiral narrative to manipulate the storyline of the IVR SG (i.e., redo from the point when an error is made) and immediate feedback to foster reflection-in-action. Finally, an experiment assessing the effectiveness of the IVR SG training system is demonstrated, leading to the main findings and a discussion of further research.

7.1.1. Storylines and narratives

Fundamentally, an IVR SG incorporates the concept of game-based learning (GBL), which covers three essential aspects: learning, play, and storylines (Göbel et al., 2010). A storyline is a game element that presents game content with structured events (Iuppa & Borst, 2012). In GBL, learners engage with storylines to play and learn. Knowledge is conveyed in context through storylines within game environments (Göbel, Rodrigues, Mehm, & Steinmetz, 2009; Starks, 2014). The use of storylines is beneficial to improving the understanding and recall of presented materials (Bransford et al., 2012). In addition, storylines promote the immersion and motivation of learners (Padilla-Zea, Gutiérrez, López-Arcos, Abad-Arranz, & Paderewski, 2014). Appealing storylines encourage learners to engage in gameplay, leading to enhanced learning effects (Habgood, Ainsworth, & Benford, 2005). As such, storylines set out a baseline in GBL, with a strong influence on other gaming characteristics, such as interactivity, challenges, and feedback (Asgari & Kaufman, 2004).

Storylines consist of a set of scenarios. The way to serve these scenarios to learners is referred to as narratives (i.e., storytelling methods) (Chatman, 1980). Narratives are essential to communicate learning content to learners (Lim et al., 2014). In principle, there are two types of narratives for GBL: an implicit narrative and an explicit narrative (Ferguson et al., 2020). In the case of implicit narratives, the storylines are not structurally outlined. The presentation of storylines depends on the exploration of learners in gaming environments. Regarding explicit narratives, storylines are clearly presented to learners. Storylines take the lead in learner activities. Implicit narratives may lead to increased recall of spatial knowledge, whereas explicit narratives may perform better on transferring factual knowledge (Ferguson et al., 2020). As such, the present study incorporates explicit narratives to the proposed IVR SG training system.

With explicit narratives, one possible way to encourage learners to follow the lead of storylines is to use an action-driven method (Feng et al., 2018). As a storyline is a set of structured scenarios, the progression from a scenario to the next scenario is driven by the actions taken by learners. In other words, in a scenario of storylines, learners may perform some activities, such as solving problems or completing tasks, in order to make progress to the next scenarios. Thus, scenarios are exposed to learners in a structured way. Eventually, the entire storyline is disclosed explicitly to learners.

7.1.2. Reflection and redo

The term “reflection” has a variety of different interpretations in the literature. A general view refers to reflection as a form of the mental process with involvement in learning and comprehension (Moon, 2013). There are two categories of reflection: reflection-in-action and reflection-on-action (Schön, 1984). Reflection-in-action means the reflection occurs while individuals are responding to a situation (Greenwood, 1993). With the effort of making sense of confronting situations, individuals reflect on their understandings and make responses accordingly (Greenwood, 1993). Reflection-in-action reshapes the perceptions of individuals and guides individuals’ follow-on actions (Greenwood, 1993). In terms of reflection-on-action, it is a retrospective process which interprets and analyzes the recalled information on undertaken practices, upon the completion of the entire learning or training. The reflection turns the contemplation of information into knowledge through a postmortem cognitive process (Greenwood, 1993). Individuals step back into previous experience and retrieve their memories with the purpose of comprehending situations and gaining knowledge.

In GBL, one effective way to facilitate reflection is using feedback (C. I. Johnson et al., 2017). Feedback is a game feature that directs learners to evaluate their performance, identify knowledge gaps, and obtain correct knowledge by receiving various forms of information (C.

I. Johnson & Priest, 2014; Shute, 2008). When learners interact with gaming environments, they generate an effect on the environments, which in turn feeds learners with information to reaffirm or adjust the learners' approaches to the confronting situations (Schön, 1984). Reflection-in-action occurs when learners receive feedback while they are responding to the game environment. The changes in the game environment impact the perspectives of learners to the surrounding situations. With feedback and reflection, learners develop new perceptions for current situations.

At the end of the reflection, it is possible for learners to adapt themselves to current situations with newly developed perceptions, enabled by reflective redo. Reflective redo allows learners to revisit a scenario and make a different response during a training experience, usually starting from the point when an error is made (Scoresby & Shelton, 2014). With the given feedback, they redo an activity that was not done appropriately, which can lead to a change in behavior and enhance their reflective thoughts to avoid the same mistake being made again (Scoresby & Shelton, 2014). In the literature of GBL, reflective redo has been studied for After Action Review (AAR) (Scoresby & Shelton, 2014). AAR is a retrospective process which provides feedback after training exercises, similar to reflection-on-action (Baird, Holland, & Deacon, 1999; Morrison & Meliza, 1999). Scoresby and Shelton (2014) demonstrate that reflective redo helps learners develop cognition on their actions with retrospective reflection (i.e., AAR) and eventually, improve learning.

In the literature, little attention has been paid to the integration of reflective redo with reflection-in-action. This study explores the possible effect of reflection-in-action with the implementation of reflective redo, using an IVR SG training system. Reflective redo is enabled by game mechanisms applied to the narrative of the IVR SG training system. More details of the design and development of the IVR SG training system are outlined in the next section.

7.2. The IVR SG training system

The proposed IVR SG training system allows trainees to experience an indoor earthquake and post-earthquake evacuation in an office building setting. Trainees are expected to apply best evacuation practice and learn suitable safety knowledge in relation to behavioral responses. In this section, the design of storylines and narratives are discussed first, which is fundamental to enable reflective redo. Following that, the use of immediate feedback is discussed, which facilitates reflection-in-action. Lastly, the development and deployment of the IVR SG training system are outlined.

7.2.1. Storylines and narratives

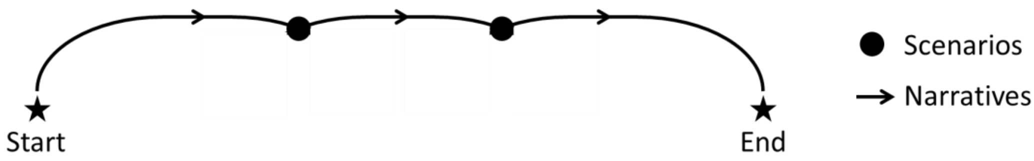
The target trainees of the IVR SG training system are adults, and the case study is the Faculty of Engineering at The University of Auckland, New Zealand (office building area). We followed the guidelines issued by the New Zealand Civil Defence and Emergency Management (NZCDEM) to identify a list of behavioural responses as the training objectives of the IVR SG training system (see Table 7-1) (New Zealand Ministry of Civil Defence & Emergency Management, 2015). The guideline includes a total of 32 recommended behavioural responses, covering an extensive range of scenarios during and after an earthquake. Note that these are recommended behavioral responses, and it does not necessarily imply that all of them must be taken into account in every single earthquake emergency. Real earthquakes are highly dynamic and unpredictable hazards, where situations change rapidly. In our case, we selected the behavioral responses focused on indoor scenarios (at work) only, which were feasible to be implemented in our proposed virtual environment (an office setting). Accordingly, several scenarios were developed to form up the storyline of the training for trainees to practice and learn knowledge in context (Starks, 2014). As we discussed in Section 7.1.1, an action-driven approach was applied to drive the storyline. The storyline was progressed once trainees had taken actions in scenarios (Feng et al., 2018). In our case, trainees had to select options to solve problems confronted in the IVR SG as the storyline goes on.

Table 7-1 Recommended behavioral responses and corresponding scenarios

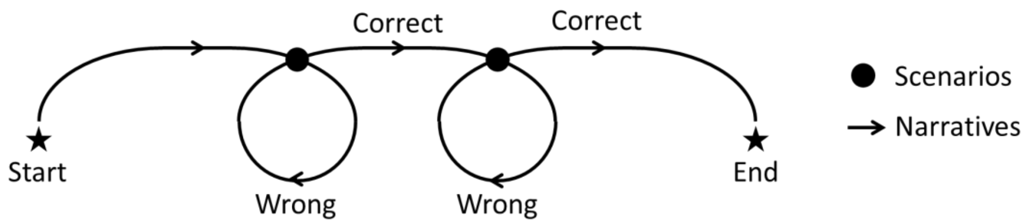
Behavioral responses	Scenarios
Drop, Cover and Hold in an earthquake	An earthquake strikes the building trainees are in. Trainees need to decide how to respond.
Do not attempt to run outside during an earthquake	Trainees can decide whether to run during the earthquake.
Take personal items when you evacuate	Before evacuation, trainees need to decide whether to take their wallet, cellphone, and laptop.
Do not use a lift when evacuate	Trainees need to decide which evacuation route to use.
Remember there may be some aftershocks	An aftershock comes when trainees are evacuating. Trainees need to decide how to respond.
Watch out for possible hazards and mitigate them if you can	An electrical appliance is emitting sparks. Trainees need to decide whether to turn off the power supply of the electrical appliance.
Stay calm. If you can, help others who may need it	A person (Non-player Character) is trapped under a table. Trainees can decide whether to offer help.

Traditionally, storylines are narrated in a linear way. A linear narrative means that storylines are disclosed from the beginning to the end with no use of flashbacks (Daniel & Rod, 2020). In the case of IVR SGs, with the integration of an action-driven approach, storylines progress from one scenario to the next after trainees take actions to solve a problem in a scenario (see Figure 7-2: a). No matter how trainees perform, trainees do not have a chance to revisit the previous scenario and adapt their previous responses. Trainees experience each scenario only once from start to end of the storyline.

In order to support the reflective redo in IVR SGs, we propose a spiral narrative to progress storylines (see Figure 7-2: b). The progress of storylines depends on the performance of trainees. Trainees make a response to a problem in a scenario. Then trainees receive immediate feedback, indicating whether their responses are appropriate. If the responses are not appropriate, trainees stay in the same scenario after receiving the feedback. Storylines are not progressed from the current scenario to the next scenario until trainees respond correctly to the problem in the current scenario. Trainees can make several attempts to solve a problem in the same scenario. As such, redo from the point when an error is made is enabled.



(a) Linear narrative



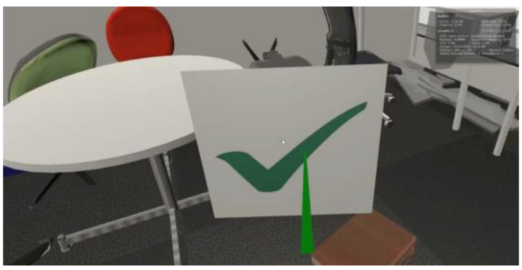
(b) Spiral narrative

Figure 7-2 Linear narrative and spiral narrative

7.2.2. Immediate feedback

In order to foster reflection, immediate feedback is given to trainees after a specific response is undertaken (Kiili, 2007). Trainees are expected to think over their previous responses based on the feedback and learn from it. In the IVR SG training system, three types of stimulation are deployed with immediate feedback: image-based feedback, audio feedback, and text-based feedback. Image-based feedback means that a green check is shown for a recommended behavioral response and a red cross for a not recommended response (see Figure 7-3: b, c). Audio feedback stands for a sound effect which is triggered simultaneously with image-based feedback, with "Ding" for a green check. Text-based feedback is presented after image-based feedback (for green checks only), which is a short text explaining the recommended behavioral response for the current problem (or scenario) (see Figure 7-3: d). Trainees receive positive feedback after a correct response and negative feedback for incorrect responses. The exposure to negative feedback may result in a negative suggestion effect, in which case misinformation could be learned by trainees (A. S. Brown et al., 1999). As such, the negative suggestion effect can jeopardize the acquisition of correct knowledge. To avoid this effect in our IVR SG training system, we propose varying degrees of stimulation for positive feedback and negative feedback. Intense stimulation is applied for positive feedback, with all the three types of stimulation triggered: when trainees choose a recommended behavioral response, a green check is popped up with the sound effect triggered at the same time, followed by a textual explanation. Whereas for negative feedback, weak stimulation is deployed with image-based feedback only: when trainees make a selection against recommended behavioral responses, a red cross is shown. The differentiation in stimulation encourages trainees to pay more attention to positive feedback, with little attention to negative feedback that only indicates a response is not recommended. In this way, the message of recommended behavioral responses is emphasized instead of misinformation (A. S. Brown et al., 1999; Butler & Roediger, 2008).

(a) A problem to solve (trainees need to



(b) Image-based feedback: a green check

find a place to take cover)

for a recommended behavioral response (taking cover under a table)

(c) Image-based feedback: a red cross for a not recommended behavioral response (trying to take cover under a table close to windows)

(d) Text-based feedback: recommended behavioral response

Figure 7-3 Immediate feedback deployed in the IVR SG training system

Another important factor facilitating effective reflection enabled by immediate feedback is the reflection time after receiving the feedback. Insufficient reflection time could lead to inadequate reflection, which can harm the acquisition of knowledge (Romano & Brna, 2001). Reflection time varies in the literature of IVR SGs, depending on the types of knowledge to be conveyed, such as procedural knowledge or factual knowledge (Krathwohl & Anderson, 2009). In an IVR SG study where participants were trained about the behavioral responses to aircraft emergencies, a 7-second reflection time was provided after each immediate feedback (Chittaro & Buttussi, 2015). Trainees could utilize this time interval to digest the feedback and trigger reflective thinking over their previous behaviours. At the same time, the IVR SG was paused temporarily without undergoing visual and acoustic simulations. In another IVR SG study where spatial knowledge was taught for evacuation, no reflection time was available for trainees when they received immediate feedback (Burigat & Chittaro, 2016). Effective training outcomes were obtained from both studies. In our case, the knowledge to be equipped by trainees is about the behavioral responses to earthquake emergencies. As such, we provided a time interval using a 10-second reflection time after text-based feedback, when trainees had completed the current scenario (i.e. they have solved the problem correctly with a recommended behavioral response). Trainees were encouraged to use the reflection time to reflect on their previous performance in response to the confronted problems by reading the text-based feedback thoroughly (Chittaro & Buttussi, 2015). The storyline of the IVR SG training system was paused during the reflection time. Once the reflection time is up, the storyline started to progress to the next scenarios.

7.2.3. The setup of the IVR SG

A virtual earthquake takes place in the IVR SG training system. An office building of the University of Auckland was selected as the training location. We followed the Building Information Modelling (BIM)-based workflow proposed by Lovreglio et al. (2018) to develop virtual environments. BIM-based workflow allows the accurate representation of building layouts and the manipulation of individual objects to permit credible earthquake simulations (Feng et al., 2018; Lovreglio et al., 2018). A basic building model defining the envelope and layout of the selected built environment was developed using Autodesk Revit (see Figure 7-4), covering walls, ceilings, floors, doors, and windows. Next, it was imported to Unity for IVR and game features development. At this stage, low-polygon models of furniture and appliances were placed in the model, allowing a fluid IVR experience. A fluid IVR experience can avoid negative symptoms and effects, such as nausea and disorientation, resulting from low frames per second (FPS) (Sharples et al., 2008).

(a) The model in Revit, without furniture and appliances (ceilings are hidden for demonstration)

(b) The final model in Unity, including furniture and appliances (ceilings are hidden for demonstration)

Figure 7-4 The models in Revit and Unity

A qualitative approach was applied to model an earthquake and provide trainees with the sense of being in an actual earthquake (Lovreglio et al., 2018). The actual performance of objects was not simulated, as the main focus of the IVR SG training system was delivering the knowledge of recommended behavioural responses, rather than structural simulations and analysis. With a qualitative approach, subjective descriptions were one of the data sources for the development of earthquake simulations (Feng et al., 2020). In our case, we referred to the New Zealand Modified Mercalli Scale (MMI) to simulate an earthquake and its damage (GeoNet, 2019). We selected the description of MMI 6: "Furniture and appliances may move on smooth surfaces, and objects fall from walls and shelves. Glassware and crockery break. Slight non-structural damage to buildings may occur" (GeoNet, 2019). The reason to use MMI 6 is that it represents a strong earthquake, enabling trainees to build a clear perceptual picture of what an earthquake looks like; however, no significant structural damage is caused at this level. Structural damage is not necessary in the IVR environment, as the message to be delivered to trainees is the recommended behavioral responses in specific scenarios, in which case, scenarios in structurally damaged buildings are not included by the New Zealand national guidelines, and as such do not help to deliver the intended training outcomes. We continued the development in Unity for earthquake simulation and building damage. We manipulated the movement, orientations, and positions of individual objects in the IVR environment, providing the visual cues of earthquakes and damage. In addition, sound effects were integrated at the same time based on the description of MMI 6.

The IVR SG training system was developed and deployed with Unity version 2018.2.14f1. A DELL PC workstation was used to run the IVR SG, which was equipped with an Intel Xeon W-2125 processor, an NVidia GeForce RTX 2080 graphics card, and 64 GB RAM. An Oculus Rift VR system enabled the IVR experience, with a head-mounted display (HMD), a remote controller, and two tracking sensors. The video output of the HMD was transmitted to an LED screen, which allowed real-time observation of the IVR experience of trainees.

7.3. Research methods

To evaluate the proposed IVR SG training system, we conducted an experimental study comparing a version of the IVR SG with redo (a spiral narrative), a version without redo (a linear narrative), and a traditional leaflet (control group). The three approaches will be referred to as Spiral narrative, Linear narrative, and Leaflet. The present study followed a pretest-posttest research design. Prior to the training via each training approach, a pretest measure on the outcome of interest was administered to trainees, followed by the same measure after training (post-test). This experimental design is good at exposing the effects of a treatment or an intervention on a group of subjects. We conducted within-groups analysis to investigate the effects of each training approaches. We also applied between-groups analysis to compare these training approaches regarding the effectiveness of knowledge gain and self-efficacy improvement. This section outlines the materials of the research and the information about trainees first. Then the measures to assess the effectiveness of the IVR SG training system are discussed. Finally, the procedure to conduct the experimental study is described.

7.3.1. Materials

The Spiral narrative and Linear narrative groups were treated with the IVR SG training system, using the software and hardware described in Section 7.2.3. The only between-group difference was the narrative mechanism, in which case Spiral narrative offered reflective redo and Linear narrative did not. Regarding Leaflet, an A4-sized paper was printed with instructions about the expected behavioral responses, as recommended in Table 7-1.

7.3.2. Trainees

Ninety-nine university students and staff (44 females and 55 males), with ages ranging from 18 to 53 years old (mean = 26.9, standard deviation = 7.74), participated in the experiment. Trainees were recruited by posters, emails, and referrals. Trainees were randomly assigned to three groups, with 33 trainees in each group.

The previous experience with earthquake drills and IVR were collected from trainees (see Table 7-2). No significant differences were revealed between groups based on Kruskal-Wallis tests (earthquake drills, $p = 0.593$; IVR, $p = 0.523$).

Table 7-2 The numbers of trainees who had previous experience with earthquake drills and IVR

Groups (N = 33)	Earthquake drills		IVR	
Spiral narrative	16	48.5%	17	51.5%
Linear narrative	20	60.6%	17	51.5%
Leaflet	17	51.5%	21	63.6%

Since the participant are the same as those in Chapter 5 (adults), the adequacy of sample size and power analysis has been discussed in Section 5.4.2.

7.3.3. Measures

The research questions focus on the effectiveness of the IVR SG training system in terms of delivering training outcomes. The training outcomes lie in the enhancement of the safety knowledge of appropriate behavioral responses and the self-efficacy in coping with earthquake emergencies (Feng et al., 2020). Following a pretest-posttest research design, a questionnaire (see Section 7.3.3.1 and 7.3.3.2) measuring the safety knowledge and self-efficacy of trainees were administered before and after the execution of the training. In addition, after training, we collected user feedback on engagement, mainly about attention (see Section 7.3.3.3). We also measured the perceptions of trainees towards the deployed narrative and feedback mechanisms in the IVR SG training system (see Section 7.3.3.4).

7.3.3.1. Safety knowledge

Trainees are expected to be able to deal with similar real-life scenarios after training (Li et al., 2017). In our case, the safety knowledge learned through training is the appropriate behavioral responses to an earthquake and post-earthquake evacuation, as recommended by national guidelines (see Table 7-1). In order to measure the acquisition of safety knowledge, a true-false knowledge test was established containing ten questions (see Table 7-3). Trainees were instructed to identify true statements only. Possible test scores ranged from 0 to 10, where trainees lost one mark if they missed a true statement or picked a false statement as a true one. Table 7-3 illustrates the statements and the correct answers to the knowledge test.

Table 7-3 Knowledge test

Statements	**True or false**
You can start running outside when an earthquake is hitting the building you are in.	False
You need to take your personal items before leaving.	True
You can take cover under any furniture during an earthquake.	False
You can start evacuation in the middle of the shaking.	False

You can use the fastest route to evacuate, like a lift.	False
Aftershocks can happen at any time after a major shock.	True
You can keep running outside if an aftershock comes.	False
You need to pay attention to hazards around you during evacuation.	True
If hazards are on your way out, you can ignore them to speed up evacuation.	False
Do not pay attention to others because your personal safety comes first.	False

7.3.3.2. Self-efficacy

Self-efficacy means the beliefs people hold in their competency to solve problems and overcome difficulties (Bandura, 1982; Bandura, 2010). High-level self-efficacy may result in a change in behavior, which leads to the improvement of performance when dealing with problems and difficulties (Bandura & Adams, 1977; Bandura, 1977; Stajkovic & Luthans, 1998). The General Self-Efficacy Scale has been suggested to measure self-efficacy in the literature, providing a list of statements (Schwarzer & Jerusalem, 2010). Based on those statements, a six-statement self-efficacy test was developed, focusing mainly on the perceptions towards earthquakes and post-earthquake evacuation:

1. "I know what to do when facing an earthquake";
2. "I can remain calm when facing an earthquake";
3. "I have the confidence to deal with an earthquake emergency";
4. "I can come up with a plan for responses to an earthquake";
5. "I can handle situations during an earthquake";
6. "I can think of a solution if I am in trouble during an earthquake."

Trainees were asked to rate their levels of agreement to each statement based on a 7-point Likert scale, with -3 for totally disagree and +3 for totally agree. The total score was calculated by finding the sum of the scores for each statement, with a higher score representing a higher level of self-efficacy (Schwarzer & Jerusalem, 2010). In our case, the possible total score ranged from -18 to 18.

7.3.3.3. Attention

We deployed a self-reported questionnaire to assess to what extend the training approaches attracted trainees' attention. The questionnaire included three statements, following the measurements applied by Burigat and Chittaro (2016):

1. "It was easy for me to concentrate on my learning";
2. "It was easy for me to stay focused on the task";
3. "I felt the training was fun."

Trainees answered the questionnaire based on a 7-point Likert scale, with -3 for totally disagree and +3 for totally agree. The total score was calculated by finding the mean of the three statements for each trainee (Burigat & Chittaro, 2016).

7.3.3.4. Ease of training

In this study, ease of training represents the ease of narratives and feedback to facilitate the learning process in the IVR SG training system. Following the measurements deployed by Chittaro and Sioni (2015), we developed a questionnaire, including four statements for trainees to answer:

1. "The training storyline helped me to learn";
2. "It was easy for me to understand the learning content";
3. "It was easy for me to learn about what to do during and after earthquakes";
4. "It was easy for me to remember what I have learned."

Trainees were asked to rate their levels of agreement to each statement based on a 7-point Likert scale, with -3 for totally disagree and +3 for totally agree. The total score was calculated by finding the mean of the four statements for each trainee (Chittaro & Sioni, 2015).

7.3.4. Procedure

This experimental study took place at the University of Auckland, New Zealand. Trainees were informed via a participation information sheet that the experiment involved a visual simulation using an IVR headset. Trainees were randomly assigned to three groups prior to participation. Upon arrival, trainees gave their consent for their participation and the collection of data for research analysis, by signing consent forms. The ethics approval (Protocol No. 016763) was granted from The University of Auckland Human Participants Ethics Committee. Trainees could withdraw their participation at any time without giving any reason. Then, trainees answered a questionnaire which covered demographic information, prior experiences with earthquake drills and IVR, a knowledge test, and a self-efficacy test.

Next, trainees in the IVR groups received an induction about using IVR as well as health and safety instructions. After that, trainees put on an IVR headset and were assisted in getting a clear view with it. Personal glasses were kept on where possible. Once the IVR session started, trainees received a tutorial to familiarize themselves with IVR environments as well as the interaction with problems and immediate feedback. The actual training took place once trainees were comfortable with the controls in the IVR.

Regarding the leaflet group, trainees were trained through reading a leaflet. Trainees were instructed to study the leaflet carefully till they fully understood the content, no matter how long it took.

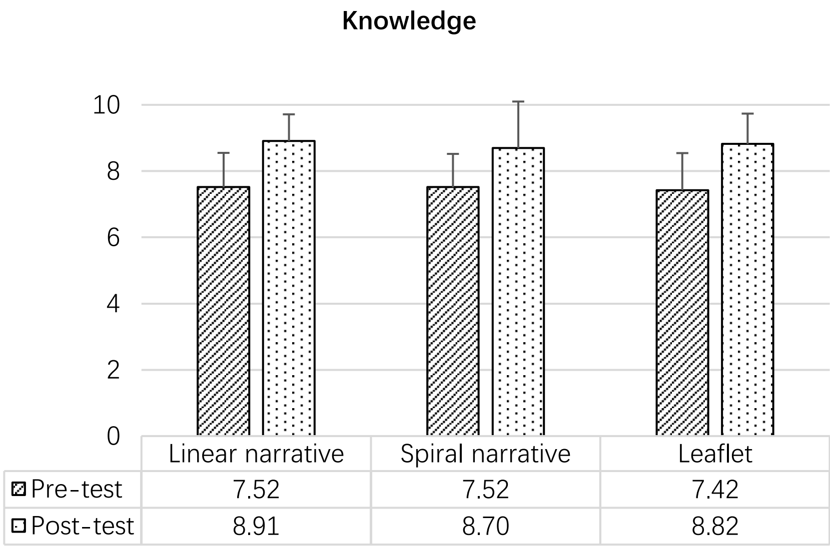
Upon the completion of training sessions, the trainees of each group answered the knowledge and self-efficacy test. Lastly, trainees were thanked, and their participation was acknowledged.

7.4. Results

The mean values of knowledge scores and self-efficacy scores are reported in Figure 7-5. Regarding the linear narrative group, the mean pre-test knowledge score was 7.52 (SD = 1.03), and the mean post-test knowledge score was 8.91 (SD = 0.80), with statistically significant improvement identified (Wilcoxon Signed Ranks Test $p < 0.001$); the mean pre-test self-efficacy score was 2.45 (SD = 5.96), and the mean post-test self-efficacy score was 9.79 (SD = 5.16), with statistically significant improvement identified (Wilcoxon Signed Ranks Test $p < 0.001$).

Regarding the spiral narrative group, the mean pre-test knowledge score was 7.52 (SD = 1.00), and the mean post-test knowledge score was 8.70 (SD = 1.40), with statistically significant improvement identified (Wilcoxon Signed Ranks Test $p < 0.001$); the mean pre-test self-efficacy score was 3.21 (SD = 8.92), and the mean post-test self-efficacy score was 10.24 (SD = 6.87), with statistically significant improvement identified (Wilcoxon Signed Ranks Test $p < 0.001$).

Regarding the leaflet group, the mean pre-test knowledge score was 7.42 (SD = 1.12), and the mean post-test knowledge score was 8.82 (SD = 0.92), with statistically significant improvement identified (Wilcoxon Signed Ranks Test $p < 0.001$); the mean pre-test self-efficacy score was 3.27 (SD = 4.94), and the mean post-test self-efficacy score was 10.55 (SD = 4.51), with statistically significant improvement identified (Wilcoxon Signed Ranks Test $p < 0.001$).



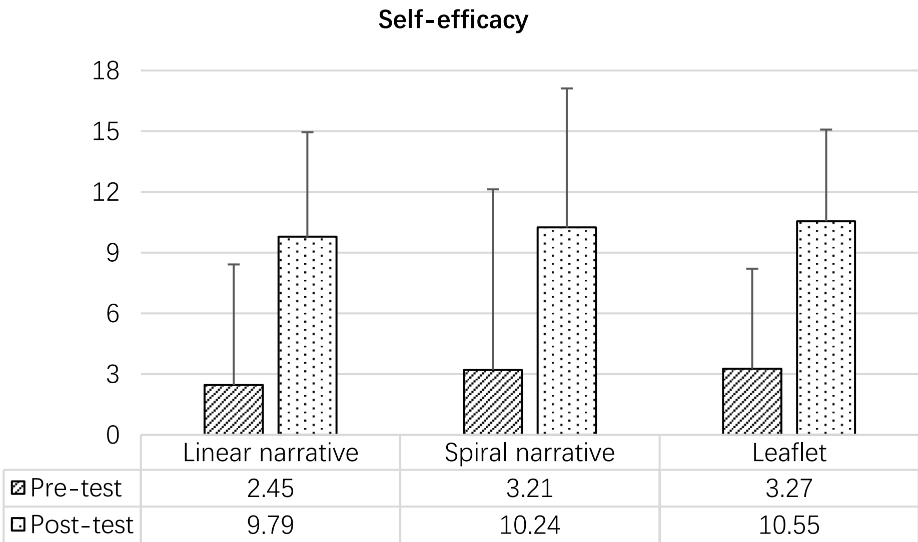


Figure 7-5 Mean pre-test and post-test knowledge and self-efficacy scores. Error bars indicate the standard deviation of the mean

ANCOVA was adopted to analyses the pre-test and post-test scores between groups (B. H. Cohen, 2013). The analysis controlled for pre-test scores as the covariate, while post-test scores were served as the dependent variable. Results revealed that there were no significant between-group differences for the post-test knowledge scores ($F(2,95) = 0.381$, $p = 0.684$, $\eta_p^2 = 0.008$) and the post-test self-efficacy scores ($F(2,95) = 0.052$, $p = 0.949$, $\eta_p^2 = 0.001$).

Bonferroni tests were applied for follow-on pairwise comparisons, as shown in Table 7-4.

Table 7-4 Between-group comparisons for knowledge and self-efficacy

Pairwise comparisons	Knowledge acquisition	Self-efficacy improvement
Leaflet vs. Linear narrative	No significant difference	No significant difference
Leaflet vs. Spiral narrative	No significant difference	No significant difference
Linear narrative vs. Spiral narrative	No significant difference	No significant difference

The mean values of attention are reported in Figure 7-6 (Linear narrative: M = 2.26, SD = 0.70; Spiral narrative: M = 2.44, SD = 0.63; Leaflet: M = 1.58, SD = 0.71). Cronbach's alphas were calculated to assess the internal consistency of multiple statements (Linear narrative: 0.868; Spiral narrative: 0.778; Leaflet: 0.768), suggesting that the three asked statements were closely related to measure attention. ANOVA was applied to determine between-group differences (B. H. Cohen, 2013). Results suggested that there was significant between-group differences for the attention scores ($F(2,96) = 14.719$, $p < 0.001$). Bonferroni tests were conducted for post-hoc pairwise comparisons, indicating a significant difference between Leaflet and Linear narrative ($p < 0.001$), a significant difference between Leaflet and Spiral narrative ($p < 0.001$), and no significant difference between Linear narrative and Spiral narrative ($p = 0.802$).

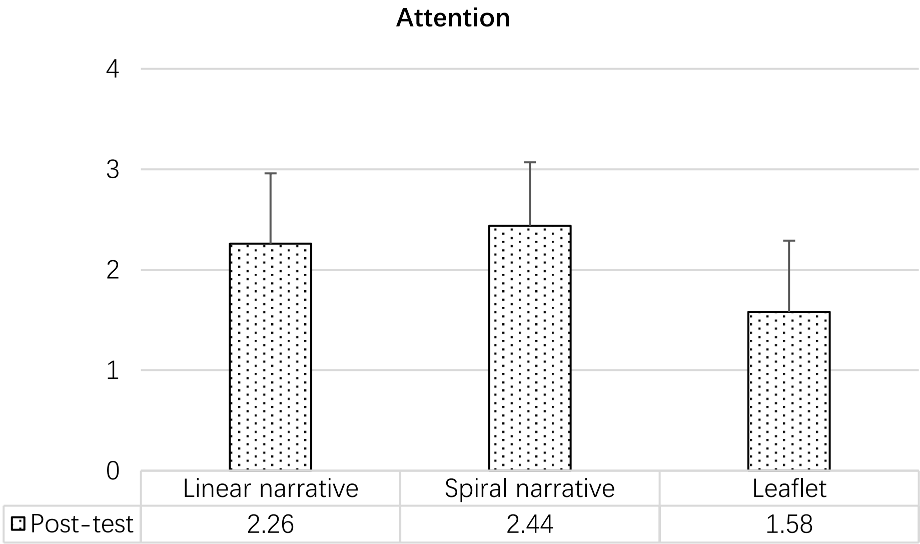


Figure 7-6 Mean attention scores. Error bars indicate the standard deviation of the mean

The results of ease of training are reported in box plots (Linear narrative: M = 2.33, SD = 0.744; Spiral narrative: M = 2.33, SD = 0.741), as shown in Figure 7-7. Cronbach's alphas were calculated to assess the internal consistency of multiple statements (Linear narrative: 0.871; Spiral narrative: 0.899), suggesting that the four asked statements were closely related to the measure on ease of training. Lastly, Kruskal-Wallis tests confirmed no significant difference between groups ($p = 0.990$).

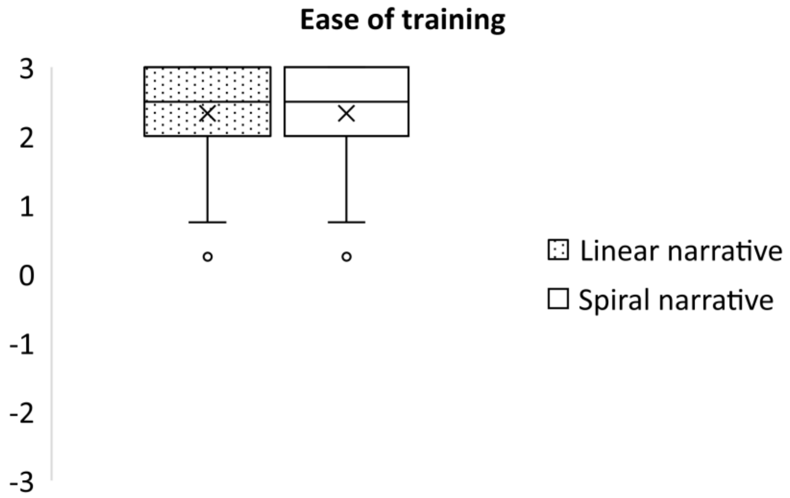


Figure 7-7 The ease of training reported by trainees

7.5. Discussion

Overall, the results point to a positive effect of the proposed IVR SG training system. However, no outperformance was observed. In this section, we discuss the obtained results.

7.5.1. Linear vs. Spiral narrative

The results of the present experiment revealed that the Linear narrative and Spiral narrative were both effective for knowledge gain and self-efficacy improvement. The similarity of Linear narrative and Spiral narrative lies in the teaching approach of immediate feedback. With immediate feedback, the trainees in both groups could develop knowledge. This finding is in line with other IVR SGs studies in the literature that immediate feedback is an effective pedagogical approach to apply in IVR SGs (Burigat & Chittaro, 2016; Feng et al., 2020).

We did not find a significant difference between the two versions of IVR SG, in terms of immediate training outcomes, attention, and the perceptions of trainees about ease of training. The manipulation of narratives did not add extra value to linear-narrated IVR SGs. It is possible that the relatively simple knowledge to be taught might weaken possible differences, given that the trainees in both groups were already knowledgeable regarding the appropriate behavioral responses for earthquakes before training (the pre-test knowledge scores from Linear and Spiral narrative groups were both 7.52 out of 10). The potential for knowledge improvement might be limited. We are uncertain how Linear and Spiral narrative will play out in a larger population with different samples.

7.5.2. IVR SG training vs. leaflet

Both Linear narrative and Spiral narrative performed better than Leaflet in terms of attracting trainees' attention, as shown in Figure 7-6. This finding further supports the notion that IVR SGs engage trainees' attention, while traditional approaches have limited ability to do so (Burigat & Chittaro, 2016; Chittaro & Buttussi, 2015). One possible reason is that IVR SGs entail trainees interacting with confronted scenarios in IVR environments, which in turn encourages trainees to stay focused on the learning materials and process (Burigat & Chittaro, 2016). Such interaction might reduce the gap between theory and practice, which is lacking in traditional approaches (Li et al., 2017).

The three tested approaches were all effective in enhancing earthquake preparedness, manifesting in significant improvement on safety knowledge and self-efficacy. However, the two versions of IVR SG did not outperform the Leaflet. This finding is consistent with other IVR SGs studies in the literature (Chittaro & Buttussi, 2015; S. Smith & Ericson, 2009), where IVR SGs are the same as traditional approaches in increasing knowledge immediately after training. However, we speculate that differences exist in a long-term effect. Chittaro and Buttussi (2015) suggest that IVR SGs could lead to better knowledge retention than traditional approaches (in their case, safety cards). In their study, after one week, trainees who were trained with safety cards suffered a significant knowledge loss, while trainees who used the

IVR SG maintained their knowledge well. One possible contributor to the retention effect of IVR SGs is the emotional arousal triggered by an engaging and emotive IVR experience, in which case memory is enhanced by emotion (Finn & Roediger III, 2011; Kensinger, 2009; Sharot, Delgado, & Phelps, 2004). Traditional approaches, such as safety cards or leaflets, are incapable of arousing intense emotion (Chittaro & Buttussi, 2015).

7.5.3. Challenges and opportunities

Based on the results, with or without repeating a scenario and redoing a response did not make a difference in the training outcomes. We measured immediate knowledge gain and self-efficacy improvement. It is possible that reflective redo might be influential in other aspects. For instance, a recent IVR SG study shows that repeated exposure to a fire emergency scenario could lessen trainees' anxiety and stress and improve wayfinding performance (Lin et al., 2019). The same effect might occur for earthquake emergency situations. As such, we speculate that trainees may develop self-efficacy and competence to reach a balance of their capability to perform activities and the difficulties of activities, which in turn facilitates a state of flow during a training experience (Admiraal et al., 2011). Flow is a state in which trainees are highly engaged in activities and are highly functional to manage activities (Czikszentmihalyi, 1990; Nakamura & Csikszentmihalyi, 2014). Flow plays an important role in GBL by keeping trainees concentrate on learning (Admiraal et al., 2011). The relationship between flow and the use of reflective redo with a spiral narrative in IVR SGs remains unclear, with a knowledge gap existing in the understanding of the impact of reflective redo on trainees' emotional states, mental workload, and cognitive activities. Future research on these topics is therefore suggested, using psychological and physiological measures.

Further to the assessment of cognitive activities, metacognition is one of the activities which has been studied in the literature with the use of reflective redo (Scoresby & Shelton, 2014).

Metacognition is referred to as "the deliberate conscious control of one's own cognitive actions" (A. L. Brown, 1980); in other words, the knowledge that people hold about their thoughts. Scoresby and Shelton (2014) investigated the metacognition of trainees who had undertaken reflective redo to get insights on the impacts of reflection on learning, with a focus on reflection-on-action. Future research can look into the metacognition associated with reflection-in-action, which is posed by a spiral narrative and reflective redo.

As well as the further investigation on the impacts and effects of reflective redo, there are opportunities to extend and escalate the use of reflective redo by improving game mechanisms. One possible way is to integrate situated learning, which suggests that learning in authentic contexts is most effective (Stein, 1998). In specific, after the initial attempts of trainees to solve problems in scenarios, their knowledge gaps are exposed. Then trainees can experience a similar scenario to solve the identical problem, with contexts (i.e., social and physical environments) being the only differentiating factor. Thus, reflective redo takes place in a different context. The repeated teaching and practicing in expansive contexts are beneficial for trainees to transferring knowledge and applying skills to new settings in the future (Engle et al., 2012; National Research Council, 2000), which is essential for earthquake safety as it is spatially unpredictable in times of need.

7.6. Conclusions

The present study contributes to the current literature on extending PBG to facilitate reflection-in-action and reflective redo. In order to do so, a spiral narrative is proposed to incorporate with immediate feedback. With the manipulation of narratives, trainees are allowed to repeat a scenario with feedback to induce reflection-in-action. An IVR SG was developed to test the game mechanisms. A linear narrative and spiral narrative were compared, with a leaflet approach being a control group. Immediate knowledge gain, self-efficacy improvement, attention, and ease of training were measured. Results support that the IVR SG training system is well aligned to PBG and superior in engagement to a traditional approach. Trainees reported that their attention was better engaged with the IVR SG training system with either narrative approach. Trainees also believed that both narrative approaches were easy for them to understand learning materials and facilitate learning processes. The results about training outcomes also suggest that a spiral narrative with immediate feedback is effective to deliver knowledge and improve self-efficacy, as well as a linear narrative and a leaflet.

The present study has several limitations, with one of them being the lack of a retention test. According to our results, reflective redo and reflection-in-action could result in a positive effect on immediate knowledge gain. As well as this, with repeated exposures to a scenario, trainees might lessen anxiety and stress, leading to improved self-efficacy, as discussed in Section 7.5.3. IVR SGs are also likely to have a long-term impact on memory. Future research can clarify the retention effect on knowledge and self-efficacy with the use of reflective redo in IVR SGs. Another limitation is that the selection of reflection time is arbitrary. Trainees were given 10 seconds after receiving text-based feedback for reflection. There is a lack in the literature about the use of reflection time in IVR SGs. We made our choice based on the literature, as discussed in Section 7.2.2. It is unclear whether a 10-second reflection time is sufficient, or too much that might interrupt a sense of presence. Lastly, the range of our sample size and characteristics is limited, with most of the trainees being knowledgeable about earthquake safety knowledge before the training. Future research can expand the experiment to other countries with different types of trainees in various settings.

This chapter investigates the pedagogical impacts of different narrative methods proposed in the customization framework (see Figure 5-1). At this stage, the effectiveness and usability of a customizable IVR SG have been discussed in Chapter 5, 6, and 7. The next chapter concludes the presented research.

Chapter 8: Conclusions

8.1. Summary of Findings

As we discussed in Section 1.2, there is a lack of research in the literature regarding IVR SG-based earthquake emergency training. In addition, the ways to deliver customizable IVR SG-based earthquake emergency training remain unclear. This research has been conducted to fill these knowledge gaps. The main findings and contributions of this research are reported in this section.

The first research objective was to identify the key factors that contribute to the development and implementation of IVR SGs in the context of emergency training. In Section 3.2, a general framework for developing and implementing IVR SG-based evacuation studies was established based on the investigation into the literature. Before the development of an IVR SG, the major work includes the recognition of outcomes (i.e., pedagogical impacts or behavioral impacts) and the selection of IVR equipment. During the development of gaming environments, a few key factors need to be considered, including teaching methods, navigation, hazards simulation, narratives, non-player characters, and sensory stimulation. In the implementation stage, outcome measures and participation experience measures were suggested to assess the usability and effectiveness of IVR SGs. This framework sets out the fundamental factors and steps for emergency studies using IVR SGs in the future.

In order to validate the proposed general framework, Chapter 4 presented a case study assessing the usability and effectiveness of an IVR SG suited to earthquakes and post-earthquake evacuation training. The proposed IVR SG was developed based on the general framework proposed in Section 3.2. The results demonstrated that the IVR SG training system improved trainees' knowledge on appropriate behavioral responses and their self-efficacy significantly. Trainees also acknowledged that the IVR SG was easy to use, and the IVR experience was engaging. The results imply that the general framework is capable of guiding the development and implementation of an IVR SG suited to earthquake emergency training.

The second research objective was to develop a framework including the atomic elements for an IVR SG-based customization framework suited to earthquake emergency training. Based on the investigation into the general framework proposed in Section 3.2 and the adaptive game-based learning framework suggested in the literature, a customization framework was proposed in Chapter 5, covering four essential dimensions: learner specifics, context, representation, and pedagogy. Pertaining to these four dimensions, six parameters were proposed for user-based customization, including trainee type, virtual environments, earthquakes and damage, storylines, storytelling methods, and teaching methods. The customization framework implies a layer-based development process, with learner specifics being the base layer and the rest of dimensions being layered atop one another to build up the customized training program (see Section 5.3.7).

The third research objective was to develop a customizable IVR SG training system to study the usability and effectiveness of the propose training framework in terms of delivering training outcomes, suited to earthquake emergency training. A case study was outlined in Section 5.4, assessing usability. A customizable IVR SG training system targeting earthquake emergency training was developed and tested with end-users. The results showed that the customization process was easy to understand and undertake, the customized virtual environments were credible, the customized storylines were suitable to trainees, the customized teaching methods facilitated learning, and the overall IVR experience was engaging. The results suggest that the proposed customization framework was capable of guiding the development of a customized IVR SG training system, with each essential dimension functioning appropriately. Chapter 6 assessed the effectiveness of different teaching methods in a customizable IVR SG-based earthquake emergency training system, including immediate feedback, post-game assessment, and prior instructions. A case study was utilized, with secondary school students being the test subjects. Results indicated that post-game assessment and prior instructions were both effective to convey knowledge and improve self-efficacy. Results implied that in the customization framework, post-game assessment and prior instructions may be the suitable teaching methods for children as trainees. Chapter 7 assessed the impacts of different storytelling methods in a customizable IVR SG-based earthquake emergency training system, covering a linear narrative and a spiral narrative. A case study showed that both linear narratives and spiral narratives were effective to facilitate learning processes, showing that trainees enhanced earthquake safety knowledge and self-efficacy significantly after training. In addition, trainees also reported that the IVR SG training systems with both narratives provided engaging training experience, and the narratives helped them comprehend learning content. This finding suggests that both narratives in the customization framework are capable of delivering functional training content and beneficial to enhancing training outcomes.

8.2. Contributions

This research offers several contributions. Firstly, this research constructs a general framework that guides the development and implementation of IVR SGs for emergency studies, which is insufficiently discussed in the current literature. This general framework contributes to existing knowledge of IVR SGs by outlining essential factors suited to IVR SGs-based training studies and behavioral analysis studies in the domain of emergencies and disasters. The empirical findings validating the general framework provides a new understanding of the applicability and effectiveness of IVR SGs-based earthquake emergency training.

Secondly, this research extends the knowledge of adaptive game-based learning by integrating

IVR SGs, leading to the establishment of a customization framework for IVR SGs suited to earthquake emergency training. The customization framework lays the groundwork for future research into customized and personalized training through IVR SGs.

Thirdly, this research contributes to the knowledge of problem-based gaming by investigating the pedagogical aspects of IVR SGs. The empirical findings advance the understanding of the effectiveness of teaching methods (i.e., immediate feedback, prior instruction, and post-game assessment) and storytelling methods (i.e., linear narratives and spiral narratives) in IVR SGs regarding immediate knowledge acquisition and self-efficacy improvement.

Fourthly, this research adds to the growing body of research that investigates the impacts of IVR SGs on children regarding earthquake emergency training. The empirical findings give insights of delivering training outcomes by IVR SGs with different teaching methods (i.e., immediate feedback, prior instruction, and post-game assessment).

Fifthly, this research has been one of the first attempts to investigate the impacts of different storytelling methods (i.e., linear narratives and spiral narratives) in IVR SGs. This research proposes a novel storytelling method to be incorporated in IVR SGs, which is a spiral narrative. The empirical findings extend the body of knowledge by investigating the training outcomes of IVR SGs with linear narratives and spiral narratives.

8.3. Limitations and Future Research

This research has several limitations, which can be addressed in future research. Firstly, the systematic literature review in Section 3.2 excluded theoretical papers that did not include follow-up case studies. The proposed general framework was established based on the empirical evidence extracted from the prototypes existing in the literature. Some other aspects and factors may be influential to the successful development and implementation of IVR SGs for emergency studies. The proposed general framework is not inclusive. Future research can extend the framework to be more inclusive and robust.

Secondly, the case study presented in Chapter 4 did not include a control group. In related works, Huang et al. (2010) and Guillén-Nieto and Aleson-Carbonell (2012) applied a similar one-group pretest-posttest experimental method as in this paper to study learning effectiveness of IVR and SGs. However, this experimental method is limited in certainty as there are threats to internal validity (Allen, 2017). It would be instructive for any future study to consider a comparative analysis between the current earthquake evacuation preparedness methods and IVR SG to examine the performance benefits. Another limitation of this case study is the lack of knowledge retention assessment. How well participants retained the newly grasped knowledge through the proposed IVR SG over time is uncertain. This is important to shift participants' behaviors in the future when they are facing actual earthquake emergencies. In a recent study, IVR SG has been suggested to outperform traditional training method (safety cards) regarding knowledge retention (Chittaro & Buttussi, 2015). One week after training, participants who were trained with safety cards suffered a significant knowledge loss, while those trained by the IVR SG retained their knowledge well. One possible way to improve knowledge retention is to induce high-level protection motivation and psychological arousal yielded by appealing experience and threat simulations (Chittaro & Sioni, 2015; Chittaro & Buttussi, 2015). Another possible way to promote retention is to introduce repetitive rehearsals with multiple environmental contexts in IVR SGs, which has been suggested to be useful to enhance memory recall (Steven M. Smith, 1982). Future research is therefore suggested to validate IVR SGs against traditional training approaches and look into the retention impact of IVR SGs.

Thirdly, the customization framework proposed in Chapter 5 faces a few challenges. One challenge is the richness of content libraries. We introduced content libraries such as an environment library and story library to offer customizable content. In this case, the possibility of customization outcomes relies on the richness of libraries. For instance, in our case study, two virtual environments and nine story modules are available for trainees. A possible solution to this is to enable the interoperability of content (Streicher & Smeddinck, 2016). Interoperability allows data to be exchanged within multiple systems. By applying this, external content can be added to content libraries, and incorporated into game settings. For instance, an original BIM model can be directly imported to the environment library for virtual environment settings, or story modules from other training programs can be "borrowed" as an inventory of the story library. Another challenge is the completeness of parameters in the four dimensions of the present framework, as we discussed in Section 5.3. The parameters are recognized based on the game-based learning model and the IVR SGs-based evacuation research framework (de Freitas & Oliver, 2006; Feng et al., 2018; van Staalduinen & de Freitas, 2011). However, those models are not specifically targeting IVR SGs-based earthquake emergency training. There might be other parameters or variables that are important for IVR SGs-based earthquake emergency training. Future studies can be conducted to complement the customization framework. Beyond that, the present framework has the potential to be extended to other emergency situations, such as tsunami, fire, or terrorist attacks. More parameters and dimensions and different settings can be incorporated into this framework. A further challenge is the modularization of virtual environments. Unlike storylines which consisted of story modules, virtual environments were stored in an environment library in the form of holistic models. Future research can explore the possibility to modularize virtual environments, where environment modules can be deployed by trainees to map out a building layout as playing with building blocks. In addition, parametric modelling can be introduced to generate adaptive virtual environments as another possible approach. A recent demonstration by Finch (www.finch3d.com) showed the possibility of creating floor plans that are adapted to the designed parameters of a site. This technology could be incorporated into the framework in the future to improve the efficiency of the customization processes. Another challenge of the customization framework is that trainees may not be competent to tune parameters to make a customized IVR SG training program that might enable an optimum behavioral response as the outcome from the training. Future research can investigate the training outcomes when trainees make customization based on their own judgement. A possible solution to this challenge is to incorporate the trainers' knowledge and input to the training process by helping with the setup of the tuning parameters in the IVR SG experience. Trainers who are knowledgeable about emergency response and training environments can intervene in the customization process to generate maximum benefit for trainees. Therefore, customization is a process including the input from trainees and trainers (Kickmeier-Rust et al., 2011). Another possible solution includes automating the customization process with Artificial Intelligence and Expert Systems. Similar to the concept of system-controlled adaptation (Streicher & Smeddinck, 2016), systems can guide trainees through the setup of tuning parameters, which can be studied in future research. Lastly, the intervention from systems (i.e., system-controlled adaptation) is currently missing. The present framework was built based on user-controlled adaptation, in which case macro adaptation is mainly handled as we discussed in Section 5.3. This leaves room for micro adaptation. System-controlled adaptation can be integrated to fill this gap. Such a hybrid approach facilitates more efficient customization and more effective training (Orji et al., 2017). System-controlled adaptation autonomously adapts systems towards better suit and support for trainees, which is a dynamically optimizing process relying on real-time monitoring and evaluating trainees' performance (Streicher & Smeddinck, 2016). In the case where trainees fail to make the best configuration for themselves through user-controlled adaptation, systems-enabled micro adaptation can give trainees adaptive pedagogical support without interrupting the flow of training experiences (Kickmeier-Rust & Albert, 2010). Possible ways to incorporate system-controlled adaptation include applying machine learning to detect and predict trainees' performance.

Fourthly, the case study presented in Chapter 6 only investigates the IVR SG training system against reading material, while other traditional earthquake training methods exist such as drills and lessons. More research using other traditional training methods is needed to build a deeper understanding of the impacts of IVR SGs. In addition, how IVR SGs work as a complement to traditional training methods for children remains unclear. Future research can explore the possibility to incorporate IVR SGs with traditional training programs, such as classroom teaching. In addition, no cognitive workload and psychological activities were measured in our study. Future research can include such measurements to bring new evidence and insights about the implementation of different mechanisms in IVR SGs. Lastly, a long-period knowledge retention test was not conducted in our study. How will children keep the knowledge obtained through different training approaches in the future? This question remains unclear and requires further investigation. The lack of a retention test also exists in the case study presented in Chapter 7. The case study in Chapter 7 did not find significant differences between IVR SGs and traditional training approaches regarding the effectiveness of delivering training outcomes. Both approaches allowed positive knowledge acquisition immediately after training. However, IVR SGs can be better than traditional approaches in terms of knowledge retention, as discussed in Section 7.6. Future research can investigate the retention impact of IVR SGs integrated with linear or spiral narratives against traditional training approaches.

References

Admiraal, W., Huizenga, J., Akkerman, S., & ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185-1194.

Alexander, D. (2012). What can we do about earthquakes? towards a systematic approach to seismic risk mitigation. *Bulletin of the New Zealand Society for Earthquake Engineering, 45*(2), 49-60.

Allen, M. (2017). *The SAGE encyclopedia of communication research methods*. Thousand Oaks, California.

Andrée, K., Nilsson, D., & Eriksson, J. (2016). Evacuation experiments in a virtual reality high-rise building: Exit choice and waiting time for evacuation elevators. *Fire and Materials, 40*(4), 554-567.

Aoki, H., Oman, C. M., & Natapoff, A. (2007). Virtual-reality-based 3D navigation training for emergency egress from spacecraft. *Aviation, Space, and Environmental Medicine, 78*(8), 774-783.

Asgari, M., & Kaufman, D. (2004). *Relationships among computer games, fantasy, and learning*. Simon Fraser University.

Auckland District Health Board. (2009). *Auckland district health board evacuation plan 2009.*

Auckland District Health Board. (2014). *The emergency preparedness and response manual.*

Backlund, P., & Hendrix, M. (2013). Educational games-are they worth the effort? A literature survey of the effectiveness of serious games. Paper presented at the *2013 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES),* 1-8.

Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., & Jin, M. (2008). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *Journal of the Learning Sciences, 17*(1), 102-141.

Baird, L., Holland, P., & Deacon, S. (1999). Learning from action: Imbedding more learning into the performance fast enough to make a difference. *Organizational Dynamics, 27*(4), 19-32.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.

Bandura, A. (2010). Self-efficacy. *The Corsini Encyclopedia of Psychology*, 1-3. Wiley Online Library.

Bandura, A., & Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research, 1*(4), 287-310.

Barreto, P., Prada, R., Santos, P. A., Ferreira, M. A., O'Neill, H., & Oliveira, C. S. (2014). Treme-treme - A serious game to teach children earthquake preparedness. *Proceedings of Videojogos*.

Bartolome, N. A., Zorrilla, A. M., & Zapirain, B. G. (2011). Can game-based therapies be trusted? is game-based education effective? A systematic review of the serious games for health and education. Paper presented at the *16th International Conference on Computer Games*.

Becker, J. S., Coomer, M. A., Potter, S. H., McBride, S. K., Lambie, E. S., Johnston, D. M., . . . Walker, A. (2016). Evaluating new zealand's "ShakeOut" national earthquake drills: A comparative analysis of the 2012 and 2015 events. Paper presented at the *Proceedings of the 2016 NZSEE Conference, Christchurch, New Zealand,* 1-3.

Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of serious games: An overview. *Advances in Human-Computer Interaction, 2013*, 1-11.

Bernardini, G., D'Orazio, M., & Quagliarini, E. (2016). Towards a "behavioural design" approach for seismic risk reduction strategies of buildings and their environment. *Safety Science, 86*, 273-294.

Bernardini, G., Quagliarini, E., & D'Orazio, M. (2016). Towards creating a combined database for earthquake pedestrians' evacuation models. *Safety Science, 82*, 77-94.

Bernhardsdottir, A. E., Musacchio, G., Ferreira, M. A., & Falsaperla, S. (2016). Informal education for disaster risk reduction. *Bulletin of Earthquake Engineering, 14*(7), 2105-2116.

Bille, R., Smith, S., Maund, K., & Brewer, G. (2014). Extending building information models into game engines. *Proceedings of the 2014 Conference on interactive entertainment*, 1-8.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). TARGET ARTICLE: Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*(2), 103-124.

Boud, D., Keogh, R., & Walker, D. (2013). *Reflection: Turning experience into learning*. Routledge.

Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C. K., & Williams, S. M. (2012). Anchored instruction: Why we need it and how technology can help. *Cognition, education, and multimedia*, 129-156, Routledge.

BRANZ seismic resilience. (2019). Measuring earthquakes. Retrieved from http://www.seismicresilience.org.nz/topics/seismic-science-and-site-influences/measuring-earthquakes/

Brewer, M. B., & Crano, W. D. (2000). Research design and issues of validity. *Handbook of Research Methods in Social and Personality Psychology,* 3-16.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry, 189*(194), 4-7.

Brown, A. S., Schilling, H. E., & Hockensmith, M. L. (1999). The negative suggestion effect: Pondering incorrect alternatives may be hazardous to your knowledge. *Journal of Educational Psychology, 91*(4), 756.

Brown, A. L. (1980). Metacognitive development and reading. *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education,* 453-481.

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

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36*(3), 604-616.

Cahill, H., Beadle, S., Mitch, J., Coffey, J., & Crofts, J. (2010). Adolescents in emergencies. *Parkville: University of Melbourne.*

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., . . . Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 61*, 27-36.

Carrozzino, M., & Bergamasco, M. (2010). Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage, 11*(4), 452-458.

Chatman, S. B. (1980). *Story and discourse: Narrative structure in fiction and film.* Cornell University Press.

Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE Transactions on Visualization and Computer Graphics, 21*(4), 529-538.

Chittaro, L., Buttussi, F., & Zangrando, N. (2014). Desktop virtual reality for emergency preparedness. *Proceedings of the 20th ACM Symposium on virtual reality software and technology*, 141-150.

Chittaro, L., & Ranon, R. (2009). Serious games for training occupants of a building in personal fire safety skills. Paper presented at the *2009 Conference in Games and Virtual Worlds for Serious Applications,* 76-83.

Chittaro, L., & Sioni, R. (2015). Serious games for emergency preparedness: Evaluation of an interactive vs. a non-interactive simulation of a terror attack. *Computers in Human Behavior, 50*, 508-519.

Clariana, R. B., Wagner, D., & Murphy, L. C. R. (2000). Applying a connectionist description of feedback timing. *Educational Technology Research and Development, 48*(3), 5-22.

Coburn, A. W., Spence, R. J., & Pomonis, A. (1992). Factors determining human casualty levels in earthquakes: Mortality prediction in building collapse. *Proceedings of First International Forum on Earthquake related Casualties.*

Codreanu, T. A., Celenza, A., & Jacobs, I. (2014). Does disaster education of teenagers translate into better survival knowledge, knowledge of skills, and adaptive behavioral change? A systematic literature review. *Prehospital and Disaster Medicine, 29*(6), 629-642.

Cohen, B. H. (2013). *Explaining psychological statistics* (4th ed.). Hoboken, New Jersey: Wiley.

Cohen, D., Sevdalis, N., Taylor, D., Kerr, K., Heys, M., Willett, K., . . . Darzi, A. (2012). Emergency preparedness in the 21st century: Training and preparation modules in virtual environments. *Resuscitation, 84*(1), 78-84.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences.* Academic press.

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education* (7th ed.). Routledge.

Connolly, T. M., Boyle, E. A., Boyle, J. M., MacArthur, E., & Hainey, T. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686.

Cosma, G., Ronchi, E., & Nilsson, D. (2016). Way-finding lighting systems for rail tunnel evacuation: A virtual reality experiment with oculus rift. *Journal of Transportation Safety & Security, 8*(sup1), 101.

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. P*roceedings of the 20th annual conference on Computer graphics and interactive techniques*, 135-142.

Czikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

Daniel, C., & Rod, M. (2020). *A dictionary of media and communication* (3rd ed.) Oxford University Press.

de Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education, 46*(3), 249-264.

Didehbani, N., Allen, T., Kandalaft, M., Krawczyk, D., & Chapman, S. (2016). Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior, 62*, 703-711.

Diehl, L. A., Lehmann, E., Souza, R. M., Alves, J. B., Esteves, R. Z., & Gordan, P. A. (2011). A serious game prototype for education of medical doctors and students on insulin management for treatment of diabetes mellitus. Paper presented at the *2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*.

Doocy, S., Daniels, A., Packer, C., Dick, A., & Kirsch, T. D. (2013). The human impact of earthquakes: A historical review of events 1980-2009 and systematic literature review. *PLoS Currents, 5*.

Duarte, E., Rebelo, F., Teles, J., & Wogalter, M. S. (2014). Behavioral compliance for dynamic versus static signs in an immersive virtual environment. *Applied Ergonomics, 45*(5), 1367-1375.

Dunwell, I., Petridis, P., Arnab, S., Protopsaltis, A., Hendrix, M., & de Freitas, S. (2011). Blended game-based learning environments: Extending a serious game into a learning content management system. Paper presented at the *2011 Third International Conference on Intelligent Networking and Collaborative Systems,* 830-835.

Engineering Village. (2020). Engineering village content coverage. Retrieved from https://www.elsevier.com/solutions/engineering-village/content

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? several proposed explanations and a research agenda for investigating them. *Educational Psychologist, 47*(3), 215-231.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education, 67*, 156-167.

Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education, 127*, 252-266.

Feng, Z., González, V. A., Ma, L., Al-Adhami, M. M. A., & Mourgues, C.Rapid 3D reconstruction of indoor environments to generate virtual reality serious games scenarios. Paper presented at the *18th International Conference on Construction Applications of Virtual Reality.*

Feng, Z., González, V. A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., . . . Cabrera-Guerrero, G. (2020). An immersive virtual reality serious game to enhance earthquake behavioral responses and post-earthquake evacuation preparedness in buildings. *Advanced Engineering Informatics, 45*, 101118.

Ferguson, C., van den Broek, Egon L, & van Oostendorp, H. (2020). On the role of interaction mode and story structure in virtual reality serious games. *Computers & Education, 143*, 103671.

Finn, B., & Roediger III, H. L. (2011). Enhancing retention through reconsolidation: Negative emotional arousal following retrieval enhances later recall. *Psychological Science, 22*(6), 781-786.

Fisher, S. S., McGreevy, M., Humphries, J., & Robinett, W. (1987). Virtual environment display system. *Proceedings of the 1986 workshop on Interactive 3D graphics*, 77-87.

Forcael, E., González, V., Orozco, F., Opazo, A., Suazo, Á, & Aránguiz, P. (2014). Application of problem-based learning to teaching the critical path method. *Journal of Professional Issues in Engineering Education and Practice, 141*(3), 04014016.

Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. *eLearning & Software for Education*, (1).

Gall, M. D., Gall, J. P., & Borg, W. R. (2006). *Educational research: An introduction* (8th ed.) Pearson.

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

Gao, Y., Gonzalez, V., & Yiu, T. W. (2017). Serious games vs traditional tools in construction safety training: A review. Paper presented at the *LC3 2017*.

García, A. A., Bobadilla, I. G., Figueroa, G. A., Ramírez, M. P., & Román, J. M. (2016). Virtual reality training system for maintenance and operation of high-voltage overhead power lines. *Virtual Reality, 20*(1), 27-40.

Gelman, A., & Stern, H. (2006). The difference between “significant” and “not significant” is not itself statistically significant. *The American Statistician, 60*(4), 328-331.

GeoNet. (2019). Earthquake intensity. Retrieved from: https://www.geonet.org.nz/earthquake/intensity

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning, 29*(3), 207-219.

Göbel, S., Rodrigues, A. d. C., Mehm, F., & Steinmetz, R. (2009). Narrative game-based learning objects for story-based digital educational games. *Narrative, 14*, 16.

Göbel, S., Wendel, V., Ritter, C., & Steinmetz, R. (2010). Personalized, adaptive digital educational games using narrative game-based learning objects. Paper presented at the *Digital Techniques and Systems 5th International Conference on E-Learning and Games, Edutainment 2010,* 438-445.

Greenwood, J. (1993). Reflective practice: A critique of the work of argyris and schön. *Journal of Advanced Nursing, 18*(8), 1183-1187.

Greitzer, F. L., Kuchar, O. A., & Huston, K. (2007). Cognitive science implications for enhancing training effectiveness in a serious gaming context. *Journal on Educational Resources in Computing (JERIC), 7*(3), 2-es.

Guillén-Nieto, V., & Aleson-Carbonell, M. (2012). Serious games and learning effectiveness: The case of it’sa deal! *Computers & Education, 58*(1), 435-448.

Gwynne, S. M., Boyce, K. E., Kuligowski, E. D., Nilsson, D., Robbins, A. P., & Lovreglio, R. (2016). Pros and cons of egress drills. Paper presented at the *Interflam 2016.*

Gwynne, S. M., Kuligowski, E. D., Boyce, K. E., Nilsson, D., Robbins, A. P., Lovreglio, R., . . . Roy-Poirier, A. (2019). Enhancing egress drills: Preparation and assessment of evacuee performance. *Fire and Materials, 43*(6), 613-631.

Habgood, M. J., Ainsworth, S. E., & Benford, S. (2005). Endogenous fantasy and learning in digital games. *Simulation & Gaming, 36*(4), 483-498.

Hansman, C. A. (2001). Context-based adult learning. *New Directions for Adult and Continuing Education, 2001*(89), 43-52.

Hardy, S., Dutz, T., Wiemeyer, J., Göbel, S., & Steinmetz, R. (2015). Framework for personalized and adaptive game-based training programs in health sport. *Multimedia Tools and Applications, 74*(14), 5289-5311.

Hartson, R., & Pyla, P. S. (2012). *The UX book: Process and guidelines for ensuring a quality user experience.* Elsevier.

Henton, M. (1996). *Adventure in the classroom: Using adventure to strengthen learning and build a community of life-long learners.* ERIC.

Hettinger, L. J., & Riccio, G. E. (1992). Visually induced motion sickness in virtual environments. *Presence: Teleoperators and Virtual Environments, 1*(3), 306-310.

Huang, H., Rauch, U., & Liaw, S. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education, 55*(3), 1171-1182.

Huber, T., Paschold, M., Hansen, C., Wunderling, T., Lang, H., & Kneist, W. (2017). New dimensions in surgical training: Immersive virtual reality laparoscopic simulation exhilarates surgical staff. *Surgical Endoscopy, 31*(11), 4472-4477.

Hwang, G., Sung, H., Hung, C., Huang, I., & Tsai, C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educational Technology Research and Development, 60*(4), 623-638.

Ip, H. H., Wong, S. W., Chan, D. F., Byrne, J., Li, C., Yuan, V. S., . . . Wong, J. Y. (2018). Enhance emotional and social adaptation skills for children with autism spectrum disorder: A virtual reality enabled approach. *Computers & Education, 117*, 1-15.

Issa, R. R. A., & Zhang, J. (2015). Collecting fire evacuation performance data using BIM-based immersive serious games for performance-based fire safety design. Paper presented at the *Computing in Civil Engineering 2015,* 612-619.

Iuppa, N., & Borst, T. (2012). *Story and simulations for serious games: Tales from the trenches.* CRC Press.

Izadkhah, Y. O., & Hosseini, M. (2005). Towards resilient communities in developing countries through education of children for disaster preparedness. *International Journal of Emergency Management, 2*(3), 138-148.

Izadkhah, Y. O., & Hosseini, M. (2007). Disaster preparedness strategy through earthquake education and training of classified target groups. Paper presented at the *Proceedings of the 2nd International Conference on Integrated Natural Disaster Management (INDM),* Tehran*.*

Jarvis, S., & de Freitas, S. (2009). Evaluation of an immersive learning programme to support triage training. Paper presented at the *2009 Conference in Games and Virtual Worlds for Serious Applications,* 117-122.

Johnson, C. W. (2006). *Using computer simulations to support a risk-based approach for hospital evacuation.* Glasgow Accident Analysis Group.

Johnson, C. I., Bailey, S. K., & van Buskirk, W. L. (2017). Designing effective feedback messages in serious games and simulations: A research review. *Instructional techniques to facilitate learning and motivation of serious games*, 119-140. Springer.

Johnson, C. I., & Priest, H. A. (2014). *The feedback principle in multimedia learning.* Cambridge University Press.

Johnson, V. A., Johnston, D. M., Ronan, K. R., & Peace, R. (2014). Evaluating children's learning of adaptive response capacities from ShakeOut, an earthquake and tsunami drill in two washington state school districts. *Journal of Homeland Security and Emergency Management, 11*(3), 347-373.

Johnson, V. A., Ronan, K. R., Johnston, D. M., & Peace, R. (2014). Implementing disaster preparedness education in new zealand primary schools. *Disaster Prevention and Management, 23*(4), 370-380.

Johnson, W., & Wu, S.Assessing aptitude for learning with a serious game for foreign language and culture. Paper presented at the *International Conference on Intelligent Tutoring Systems,* 520-529.

Jones, L. M. (1994). Foreshocks, aftershocks, and earthquake probabilities: Accounting for the landers earthquake. *Bulletin of the Seismological Society of America, 84*(3), 892-899.

Kasanen, E., Lukka, K., & Siitonen, A. (1993). The constructive approach in management accounting research. *Journal of Management Accounting Research, 5*, 243.

Kelley, C. R. (1969). What is adaptive training? *Human Factors, 11*(6), 547-556.

Kenny, C. (2009). *Why do people die in earthquakes? the costs, benefits and institutions of disaster risk reduction in developing countries.* The World Bank.

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review, 1*(2), 99-113.

Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine, 96*(3), 118-121.

Kickmeier-Rust, M. D., & Albert, D. (2010). Micro-adaptivity: Protecting immersion in didactically adaptive digital educational games. *Journal of Computer Assisted Learning, 26*(2), 95-105.

Kickmeier-Rust, M. D., Augustin, T., & Albert, D. (2011). Personalized storytelling for educational computer games. Paper presented at the *International Conference on Serious Games Development and Applications,* 13-22.

Kiili, K. (2007). Foundation for problem-based gaming. *British Journal of Educational Technology, 38*(3), 394-404.

Kinateder, M., Gromer, D., Gast, P., Buld, S., Müller, M., Jost, M., . . . Pauli, P. (2015). The effect of dangerous goods transporters on hazard perception and evacuation behavior – A virtual reality experiment on tunnel emergencies. *Fire Safety Journal, 78*, 24-30.

Kinateder, M., Müller, M., Jost, M., Mühlberger, A., & Pauli, P. (2014). Social influence in a virtual tunnel fire – influence of conflicting information on evacuation behavior. *Applied Ergonomics, 45*(6), 1649-1659.

Kinateder, M., Ronchi, E., Gromer, D., Müller, M., Jost, M., Nehfischer, M., . . . Pauli, P. (2014). Social influence on route choice in a virtual reality tunnel fire. *Transportation Research Part F: Psychology and Behaviour, 26*(A), 116-125.

Kinateder, M., Ronchi, E., Nilsson, D., Kobes, M., Muller, M., Pauli, P., & Muhlberger, A. (2014). Virtual reality for fire evacuation research. *2014 Federated Conference on Computer Science and Information Systems*, 313-321.

Kobes, M., Helsloot, I., de Vries, B., & Post, J. G. (2010). Building safety and human behaviour in fire: A literature review. *Fire Safety Journal, 45*(1), 1-11.

Krathwohl, D. R., & Anderson, L. W. (2009). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives.* Longman.

Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: Immersion aids recall. *Virtual Reality, 23*(1), 1-15.

Kuhn, D., & Pease, M. (2006). Do children and adults learn differently? *Journal of Cognition and Development, 7*(3), 279-293.

Kulhavy, R. W., & Anderson, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology, 63*(5), 505.

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*(1), 79-97.

Lane, H. C., & Yi, S. (2017). Playing with virtual blocks: Minecraft as a learning environment for practice and research. *Cognitive development in digital contexts*, 145-166. Academic Press.

LaValle, S. M. (2016). *Virtual reality.* Cambridge University Press.

Li, C., Liang, W., Quigley, C., Zhao, Y., & Yu, L. (2017). Earthquake safety training through virtual drills. *IEEE Transactions on Visualization and Computer Graphics, 23*(4), 1275-1284.

Lim, T., Louchart, S., Suttie, N., Hauge, J. B., Stanescu, I. A., Ortiz, I. M., . . . Earp, J. (2014). Narrative serious game mechanics (NSGM)–insights into the narrative-pedagogical mechanism. Paper presented at the *International Conference on Serious Games,* 23-34.

Lin, J., Cao, L., & Li, N. (2019). Assessing the influence of repeated exposures and mental stress on human wayfinding performance in indoor environments using virtual reality technology. *Advanced Engineering Informatics, 39*, 53-61.

Lin, J., Zhu, R., Li, N., & Becerik-Gerber, B. (2020). Do people follow the crowd in building emergency evacuation? A cross-cultural immersive virtual reality-based study. *Advanced Engineering Informatics, 43*, 101040.

Lindell, M. K., Prater, C. S., Wu, H. C., Huang, S., Johnston, D. M., Becker, J. S., & Shiroshita, H. (2016). Immediate behavioural responses to earthquakes in christchurch, new zealand, and hitachi, japan. *Disasters, 40*(1), 85-111.

Lopes, R., & Bidarra, R. (2011a). Adaptivity challenges in games and simulations: A survey. *IEEE Transactions on Computational Intelligence and AI in Games, 3*(2), 85-99.

Lopes, R., & Bidarra, R. (2011b). A semantic generation framework for enabling adaptive game worlds. Paper presented at the *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology,* 1-8.

Lori, P. (2008). Children and disasters: Understanding vulnerability, developing capacities, and promoting resilience — an introduction. *Children Youth and Environments, 18*(1), 1-29.

Lovreglio, R., González, V. A., Feng, Z., Amor, R., Spearpoint, M., Thomas, J., . . . Sacks, R. (2018). Prototyping virtual reality serious games for building earthquake preparedness: The auckland city hospital case study. *Advanced Engineering Informatics, 38*(2018), 670-682.

Lovreglio, R., Gonzalez, V., Amor, R., Spearpoint, M., Thomas, J., Trotter, M., & Sacks, R. (2017). The need for enhancing earthquake evacuee safety by using virtual reality serious games. Paper presented at *LC3 2017*.

Lukka, K. (2003). The constructive research approach *Case Study Research in Logistics, publications of the Turku School of Economics and Business Administration, Series B, 1*(2003), 83-101.

Mahdavifar, M. R., Izadkhah, Y. O., & Heshmati, V. (2009). Appropriate and correct reactions during earthquakes:" Drop, cover and hold on" or" triangle of life". *Journal of Seismology and Earthquake Engineering, 11*(1), 41-48.

Mathan, S. A., & Koedinger, K. R. (2002). An empirical assessment of comprehension fostering features in an intelligent tutoring system. Paper presented at the *International Conference on Intelligent Tutoring Systems,* 330-343.

Mathieu Muratet, Patrice Torguet, Jean-Pierre Jessel, & Fabienne Viallet. (2009). Towards a serious game to help students learn computer programming. *International Journal of Computer Games Technology, 2009*, 1-12.

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition, 136*, 156-165.

Mayer, I., Wolff, A., & Wenzler, I. (2013). Learning efficacy of the 'hazard recognition' serious game. Paper presented at the *International Conference on Serious Games Development and Applications,* 118-129.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research, 42*(3), 241-265.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology, 35*(4), 229-241.

McCrudden, M. T., Schraw, G., & Hartley, K. (2006). The effect of general relevance instructions on shallow and deeper learning and reading time. *The Journal of Experimental Education, 74*(4), 291-310.

McCrudden, M. T., Schraw, G., & Kambe, G. (2005). The effect of relevance instructions on reading time and learning. *Journal of Educational Psychology, 97*(1), 88.

McSaveney, E. (2017). *Earthquakes* Ministry for Culture and Heritage Te Manatu Taonga. Retrieved from https://teara.govt.nz/en/12079

Mellor, D., & Moore, K. A. (2013). The use of likert scales with children. *Journal of Pediatric Psychology, 39*(3), 369-379.

Meng, F., & Zhang, W. (2014). Way-finding during a fire emergency: An experimental study in a virtual environment. *Ergonomics, 57*(6), 816-827.

Metcalfe, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children’s and adults’ vocabulary learning. *Memory & Cognition, 37*(8), 1077-1087.

Metello, M. G., Casanova, M. A., & Carvalho, M. T. M. (2008). Using serious game techniques to simulate emergency situations. Paper presented at the *X Brazilian Symposium on Geoinformatics,* 121-182.

Michael, D., & Chen, S. (2006). *Serious games: Games that educate, train and inform*. Mason, OH: Course Technology.

Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification, 5*(2), 181-204.

Ministry of Civil Defence & Emergency Management. (2019). Earthquakes. Retrieved from https://getready.govt.nz/emergency/earthquakes/

Ministry of Civil Defence and Emergency Management. (2019). What's the plan stan? learn about earthquakes. Retrieved from https://getready.govt.nz/prepared/school/whats-the-plan-stan/learn-about-earthquakes/

Ministry of Education. (2019). Preparing for emergencies, traumatic incidents, evacuations and lockdowns. Retrieved from https://www.education.govt.nz/school/health-safety-and-wellbeing/emergencies-and-traumatic-incidents/preparing-for-emergencies-and-traumatic-incidents/

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097.

Moon, J. A. (2013). *Reflection in learning and professional development: Theory and practice.* Routledge.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*(1-2), 99-113.

Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour Research and Therapy, 74*, 18-24.

Morrison, J. E., & Meliza, L. L. (1999). Foundations of the after action review process. *United States Amy Research Institute for the Behavioral and Social Sciences.*

Morrongiello, B. A., Corbett, M., Milanovic, M., Pyne, S., & Vierich, R. (2015). Innovations in using virtual reality to study how children cross streets in traffic: Evidence for evasive action skills. *Injury Prevention, 21*(4), 266-270.

Mutch, C. (2014). The role of schools in disaster preparedness, response and recovery: What can we learn from the literature? *Pastoral Care in Education, 32*(1), 5-22.

Mutch, C. (2018). The role of schools in helping communities copes with earthquake disasters: The case of the 2010–2011 new zealand earthquakes. *Environmental Hazards, 17*(4), 331-351.

Mutch, C., & Marlowe, J. (2013). “Sailing through a river of emotions”: Capturing children's earthquake stories. *Disaster Prevention and Management, 22*(5), 445-455.

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. *Flow and the foundations of positive psychology*, 239-263. Springer.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition.* National Academies Press.

New Zealand Ministry of Civil Defence & Emergency Management. (2015). *Working from the same page consistent messages for CDEM - part B: Hazard specific information - earthquake*

New Zealand Police. (2011). *List of deceased - christchurch earthquake.* Retrieved from http://www.police.govt.nz/major-events/previous-major-events/christchurch-earthquake/list-deceased

New Zealand ShakeOut. (2018). New zealand ShakeOut - get ready to ShakeOut! Retrieved from https://www.shakeout.govt.nz/

Oculus. (2020). Oculus. Retrieved from www.oculus.com

Orji, R., Oyibo, K., & Tondello, G. F. (2017). A comparison of system-controlled and user-controlled personalization approaches. Paper presented at the *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization,* 413-418.

Padilla-Zea, N., Gutiérrez, F. L., López-Arcos, J. R., Abad-Arranz, A., & Paderewski, P. (2014). Modeling storytelling to be used in educational video games. *Computers in Human Behavior, 31*, 461-474.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1-12.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45-77.

Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes/no and wh-questions. *Law and Human Behavior, 23*(5), 539-555.

Porter, K., Kennedy, R., & Bachman, R. (2007). Creating fragility functions for performance-based earthquake engineering. *Earthquake Spectra, 23*(2), 471-489.

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science, 23*(5/6), 405-431.

Ragavan, S. V., Ponnambalam, S. G., & Sumero, C. (2011). Waypoint-based path planner for mobile robot navigation using PSO and GA-AIS. Paper presented at the *Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE,* 756-760.

Ramirez, M., Kubicek, K., Peek-Asa, C., & Wong, M. (2009). Accountability and assessment of emergency drill performance at schools. *Family & Community Health, 32*(2), 105-114.

Regan, E. C., & Price, K. R. (1994). The frequency of occurrence and severity of side-effects of immersion virtual reality. *Aviation, Space, and Environmental Medicine, 65*(6), 527.

Rhebergen, M. D., Hulshof, C. T., Lenderink, A. F., & van Dijk, F. J. (2010). An online network tool for quality information to answer questions about occupational safety and health: Usability and applicability. *BMC Medical Informatics and Decision Making, 10*(1), 63.

Rice, J. (2007). Assessing higher order thinking in video games. *Journal of Technology and Teacher Education, 15*(1), 87-100.

Riecke, B. E., Bodenheimer, B., McNamara, T. P., Williams, B., Peng, P., & Feuereissen Daniel. (2010). Do we need to walk for effective virtual reality navigation? physical rotations alone may suffice. Paper presented at the *International Conference on Spatial Cognition,* 234-247.

Romano, D. M., & Brna, P. (2001). Presence and reflection in training: Support for learning to improve quality decision-making skills under time limitations. *Cyber Psychology & Behavior, 4*(2), 265-277.

Ronan, K. R., Alisic, E., Towers, B., Johnson, V. A., & Johnston, D. M. (2015). Disaster preparedness for children and families: A critical review. *Current Psychiatry Reports, 17*(7), 58.

Ronchi, E., Kinateder, M., Müller, M., Jost, M., Nehfischer, M., Pauli, P., & Mühlberger, A. (2015). Evacuation travel paths in virtual reality experiments for tunnel safety analysis. *Fire Safety Journal, 71*, 257-267.

Ronchi, E., Nilsson, D., Kojić, S., Eriksson, J., Lovreglio, R., Modig, H., & Walter, A. (2016). A virtual reality experiment on flashing lights at emergency exit portals for road tunnel evacuation. *Fire Technology, 52*(3), 623-647.

Rüppel, U., & Schatz, K. (2011). Designing a BIM-based serious game for fire safety evacuation simulations. *Advanced Engineering Informatics, 25*(4), 600-611.

Sacfung, A., Sookhanaphibarn, K., & Choensawat, W. (2014). Serious game for fire safety evacuation plan. *Advanced Materials Research, 931-932*, 583-587.

Salkind, N. J. (2010). *Encyclopedia of research design.* Sage.

Sapien, R. E., & Allen, A. (2001). Emergency preparation in schools: A snapshot of a rural state. *Pediatric Emergency Care, 17*(5), 329-333.

Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J. (2007). Distinguishing between games and simulations: A systematic review. *Educational Technology & Society, 10*(3), 247-256.

Schön, D. A. (1984). *The reflective practitioner: How professionals think in action.* Basic books.

Schonauer, C., Pintaric, T., Kaufmann, H., Jansen-Kosterink, S., & Vollenbroek-Hutten, M. (2011). Chronic pain rehabilitation with a serious game using multimodal input. Paper presented at the *2011 International Conference on Virtual Rehabilitation*, 1-8.

Schroth, M. L. (1992). The effects of delay of feedback on a delayed concept formation transfer task. *Contemporary Educational Psychology, 17*(1), 78-82.

Schwarzer, R., & Jerusalem, M. (2010). The general self-efficacy scale (GSE). *Anxiety, Stress, and Coping, 12*, 329-345.

Schwebel, D. C., Combs, T., Rodriguez, D., Severson, J., & Sisiopiku, V. (2016). Community-based pedestrian safety training in virtual reality: A pragmatic trial. *Accident Analysis & Prevention, 86*, 9-15.

Scopus. (2020). Scopus content coverage guide. Retrieved from https://www.elsevier.com/solutions/scopus/how-scopus-works/content

Scoresby, J., & Shelton, B. E. (2014). Reflective redo from the point of error: Implications for after action review. *Simulation & Gaming, 45*(4-5), 666-696.

Sharma, S., Jerripothula, S., Mackey, S., & Soumare, O. (2014). Immersive virtual reality environment of a subway evacuation on a cloud for disaster preparedness and response training. Paper presented at the *2014 IEEE Symposium on Computational Intelligence for Human-like Intelligence (CIHLI)*, 1-6.

Sharot, T., Delgado, M. R., & Phelps, E. A. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience, 7*(12), 1376.

Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays, 29*(2), 58-69.

Shaw, R., Kobayashi, K. S. H., & Kobayashi, M. (2004). Linking experience, education, perception and earthquake preparedness. *Disaster Prevention and Management: An International Journal, 13*(1), 39-49.

Shendarkar, A., Vasudevan, K., Lee, S., & Son, Y. (2008). Crowd simulation for emergency response using BDI agents based on immersive virtual reality. *Simulation Modelling Practice and Theory, 16*(9), 1415-1429.

Sherman, W. R., & Craig, A. B. (2018). *Understanding virtual reality: Interface, application, and design*. San Francisco, CA: Morgan Kaufmann.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Silva, J. F., Almeida, J. E., Rossetti, R. J. F., & Coelho, A. L. (2013). A serious game for EVAcuation training. Paper presented at the *2013 IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH)*, 1-6.

Sliney, A., & Murphy, D. (2008). JDoc: A serious game for medical learning. *First International Conference on Advances in Computer-Human Interaction*, 131-136.

Smith, S., & Ericson, E. (2009). Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual Reality, 13*(2), 87-99.

Smith, S. M. (1982). Enhancement of recall using multiple environmental contexts during learning. *Memory & Cognition, 10*(5), 405-412.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240.

Starks, K. (2014). Cognitive behavioral game design: A unified model for designing serious games. *Frontiers in Psychology, 5*, 28.

Stein, D. (1998). *Situated learning in adult education*, 195-202. ERIC Clearinghouse on Adult, Career, and Vocational Education, Center on Education and Training for Employment, College of Education, the Ohio State University.

Streicher, A., & Smeddinck, J. D. (2016). Personalized and adaptive serious games. *Entertainment computing and serious games*, 332-377. Springer.

Stuart-Black, S. (2015). *Drop, cover and hold is still the right action to take.* New Zealand Ministry of Civil Defence & Emergency Management.

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games - an overview*. University of Skövde, School of Humanities and Informatics.

Tanes, Z., & Cho, H. (2013). Goal setting outcomes: Examining the role of goal interaction in influencing the experience and learning outcomes of video game play for earthquake preparedness. *Computers in Human Behavior, 29*(3), 858-869.

Tatebe, J., & Mutch, C. (2015). Perspectives on education, children and young people in disaster risk reduction. *International Journal of Disaster Risk Reduction, 14*, 108-114.

Tekeli-Yeşil, S., Dedeoğlu, N., Tanner, M., Braun-Fahrlaender, C., & Obrist, B. (2010). Individual preparedness and mitigation actions for a predicted earthquake in istanbul. *Disasters, 34*(4), 910-930.

Teng, C. (2010). Customization, immersion satisfaction, and online gamer loyalty. *Computers in Human Behavior, 26*(6), 1547-1554.

The New Zealand Parliament. (2018). Fire and emergency new zealand (fire safety, evacuation procedures, and evacuation schemes) regulations 2018. Retrieved from http://www.legislation.govt.nz/regulation/public/2018/0096/18.0/LMS46332.html

Thomé, A. M. T., Scavarda, L. F., & Scavarda, A. J. (2016). Conducting systematic literature review in operations management. *Production Planning & Control, 27*(5), 408-420.

United States Geological Survey. (2018). Earthquake facts. Retrieved from https://earthquake.usgs.gov/learn/facts.php

van der Kleij, Fabienne M, Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475-511.

van Staalduinen, J., & de Freitas, S. (2011). A game-based learning framework: Linking game design and learning. *Learning to Play: Exploring the Future of Education with Video Games, 53*, 29.

Vinten, G. (1995). Open versus closed questions–An open issue. *Management Decision, 33*(4), 27-31.

Virvou, M., & Katsionis, G. (2008). *On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE.* Elsevier.

Wabo, N. C., Örtenwall, P., & Khorram-Manesh, A. (2012). Hospital evacuation; planning, assessment, performance and evaluation. *Journal of Acute Disease, 1*(1), 58-64.

Walker, A., & Shelton, B. E. (2008). Problem-based educational games: Connections, prescriptions, and assessment. *Journal of Interactive Learning Research, 19*(4), 663-684.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on evaluation and assessment in software engineering*, 1-10.

Wouters, P., van der Spek, Eric D., & van Oostendorp, H. (2009). Current practices in serious game research: A review from a learning outcomes perspective. *Games-based learning advancements for multi-sensory human computer interfaces: Techniques and effective practices*, 232-250. IGI Global.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, Eric D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249-265.

Yang, X., & Wu, Z. (2013). Civilian monitoring video records for earthquake intensity: A potentially unbiased online information source of macro-seismology. *Natural Hazards, 65*(3), 1765-1781.

Yang, X., Wu, Z., & Li, Y. (2011). Difference between real-life escape panic and mimic exercises in simulated situation with implications to the statistical physics models of emergency evacuation: The 2008 wenchuan earthquake. *Physica A: Statistical Mechanics and its Applications, 390*(12), 2375-2380.

Ye, L., Qu, Z., Lu, X., & Feng, P. (2008). Collapse prevention of building structures: A lesson from the wenchuan earthquake. *Journal of Building Structures, 4*(007), 1-9.

Yu, K., Fan, S., & Lin, K. (2015). Enhancing students' problem-solving skills through context-based learning. *International Journal of Science and Mathematics Education, 13*(6), 1377-1401.

Zhou, Z., Chang, J. S., Pan, J., & Whittinghill, D. (2016). Alternate reality game for emergency response training: A review of research. *Journal of Interactive Learning Research, 27*(1), 77.

Zou, H., Li, N., & Cao, L. (2016). Immersive virtual environments for investigating building emergency evacuation behaviors: A feasibility study. *Proceedings of the International Symposium on Automation and Robotics in Construction, 33,* 1-8.

Appendices

Appendix A: Ethics Approval

Research Office
Post-Award Support Services

The University of Auckland
Private Bag 92019
Auckland, New Zealand

Level 10, 49 Symonds Street
Telephone: 64 9 373 7599
Extension: 83711
Facsimile: 64 9 373 7432
ro-ethics@auckland.ac.nz

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE (UAHPEC)

18-Jan-2018

MEMORANDUM TO:

Dr Vicente Gonzalez
Civil and Environmental Eng

Re: Request for change of Ethics Approval (Our Ref. 016763): Amendments Approved

The Committee considered your request for change for your study entitled **Building Quake and People – A Serious Game Platform for Informing Life Saving Strategies** and approval was granted for the following amendments on 18-Jan-2018.

The Committee approved the following amendments:

1. To add Ormiston Junior College and Auckland City Council as research sites.
2. To add students aged 11-16 years as participants.
3. To add a haptic system to the earthquake simulation.
4. To add Maximilien Pottiez, Skar Adhikari, and Carol Mutch to the research personnel.

The expiry date for this approval is 06-Apr-2019.

If the project changes significantly you are required to resubmit a new application to the Committee for further consideration.

In order that an up-to-date record can be maintained, it would be appreciated if you could notify the Committee once your study is completed.

The Chair and the members of the Committee would be happy to discuss general matters relating to ethics approvals. If you wish to do so, please contact the UAHPEC Ethics Administrators at ro-ethics@auckland.ac.nz in the first instance.

Please quote reference number **016763** on all communications with the UAHPEC regarding this application.

(This is a computer generated letter. No signature required.)

UAHPEC Administrators
University of Auckland Human Participants Ethics Committee

c.c. Head of Department / School, Civil and Environmental Eng
Assoc Prof Carol Mutch
Dr Ruggiero Lovreglio
Mr Jiamou Liu
Prof Robert Amor
Prof Pierre Quenneville
Assoc Prof Ashvin Thambyah
Zhenan Feng

Appendix B: Participant Information Sheet

Department of Civil and Environmental Engineering
Engineering Building
20 Symonds Street
Auckland 1142
New Zealand
Telephone 64 9 373 7599 ext 88166
Facsimile 64 9 373 7462

The University of Auckland
Private Bag 92019
Auckland 1142
New Zealand

PARTICIPANT INFORMATION SHEET

Research Project Title	Building Quake and People – A Serious Game Platform for Informing Life-Saving Strategies (Ref. 016763)
Researchers	Zhenan Feng, Civil and Environmental Engineering Department
Supervisor	Dr Vicente Gonzalez, Civil and Environmental Engineering Department
Co-Supervisor	Prof Robert Amor, Computer Science Department Assoc. Prof Carol Mutch, Education and Social Work Department

Purpose of this Participant Information Sheet (PIS)

The purpose of this Participant Information Sheet (PIS) is to invite you to participate in the research data collection for this research project. This project aims to the mitigation of the impact of natural hazards by proposing the development of a computer-based modelling framework using Serious Games (SG), Virtual Reality (VR), Building Information Modelling (BIM) and Agent-Based Models (ABM) able to assess occupants' behaviour in buildings in the event of an earthquake. You will be exposed to a virtual reality environment within a serious game framework (similar to video games) where you have to make decisions during a quake evacuation process in a building. Your reactions and behavior will be observed, and interviews will be used to characterize your behavior. The objectives of the interviews which will be informal are to gauge the nature of the data collected through your professional feedback. The interviews will be developed in the facilities of the organizations involved in this research.

Research Background

This research tries to comprehensively understand and predict human behavior during the evacuation of a building in the event of an earthquake. To do so, this project aims to develop a Serious Game (similar to video games, but with a "serious" purpose) for building owners, designers, regulators and emergency managers to use to model and simulate how occupants will behave during and after an earthquake in a particular building. Thus, the system "people-building-quake" (e.g., people's behavior and motion throughout a damaged structure in the event of an earthquake) will develop more robust, effective and reliable evacuation strategies. The Serious Game consists of some components to simulate a realistic quake and a building evacuation process: 1) Building Information Model, 2) Agent-Based Model, and 3) Virtual Reality Environment. It is argued that while post-earthquake evacuation can be perceived as a quite controlled process, more research is required in that respect to a) better understand human behaviour in the event of an earthquake, and b) enhance current evacuation practices as earthquake damage in a building can be a very dynamic and unpredictable process, blocking predefined evacuation routes or exits and damaging a wide range of structural and non-structural components in a building, so alternative and adaptive plans (e.g., dynamic signalising) can take place to respond to earthquake damage, representing what may be best practice.

Participation and withdrawal

Considering your experience in evacuation matters, you could learn evacuation knowledge and provide some information regarding your experience being exposed to the virtual environment of the serious game, in which a building subject to the effects of a quake is simulated. Participation in this study is entirely voluntary, and you still have the right to whether or not participate. If you decide to participate in this research, you also have the right to withdraw from participation at any time, and without any explanation.

Data Collection & Data Management

This stage will be undertaken in three stages. The first stage involves the collection of Building Information Modelling (BIM) data of the participants' building and a preliminary understanding of evacuation patterns. This will take place between March and October 2018. To collect BIM data, 2D and 3D architectural and engineering drawings from experiment buildings will be collected with the purpose to understand the geometry of non-structural, structural and architectural elements so that the Director Facilities Manager will assist with it. Also, information on the furniture and equipment available in the experiment buildings premises that can prevent an efficient evacuation process will be collected for further modelling as 3D objects.

You are asked to participate in a Serious Game. Thus, you will "play" the Serious Game using a virtual reality head-mounted display and game controller to make decisions and move within the Serious Game environment. The Serious Game development and application will be done between May 2018 and May 2019. The pilot test using the Serious Game will be done between June-October 2019. Your Serious Game session will take something in between 15-30 min. You will take part in one Serious Game session, which involves two stages. The first stage will be a mock round to familiarize with the technology basically. The second stage will be used to play the Serious Game itself. During this session (second stage), you will run through an earthquake scenario. The number of the scenarios will depend on the variables selected for manipulation, but will not exceed six. The location of the observation can be either at The University of Auckland or the specific location of the organization to which the participant belongs. Data will be video recorded and reviewed your virtual evacuation behaviors following the simulated earthquake. A coding scheme will then be iteratively developed from these virtual observations. This research is exploratory, and the codes and categories of behavior will be developed from view the virtual footage. Timing information will be collected on how long participants wait in place before evacuating and how long you engage in the activities identified in the coding scheme. Route selection will also be identified and collected.

Before and after the Serious Game scenario, there will be a questionnaire which is an essential step for the completion of this research. You will be asked about personal information, general questions about your decision making when facing risks in a building during a quake, strategies to avoid risk and other pedestrians and your overall perception and experience playing the serious game.

You are allowed to withdraw the questionnaires at any time during the session, without the need to provide any reason. Answers and data from the questionnaires will be analyzed and transferred to a draft survey information sheet in electronic format. You are allowed to review and withdraw the data provided after undertaking the questionnaire session. The collected data from the questionnaires will be kept for at least six years at the University of Auckland. All data collected will be stored in an electronic file on password protected computers. Data might be used in conferences, academic publications or presentations. However, organizations, individuals or individual responses will not be identified in any of these. All reports/results will be based on the overall results of the research.

You will be able to withdraw your Serious Game observation data should you no longer wish to participate, up to the point of the data analysis. Should a participant wish to withdraw, they will need to notify the researchers before data analysis (one week following the completion of his/her SG session).

There is some chance or motion sickness and vertigo with use of a virtual reality head-mounted display. You can withdraw at any point if you feel unwell. On top of that, all the measures related to an appropriate calibration of the virtual reality head-mounted display and recommendations for its appropriate use will be in place and communicated to the participants.

Upon request, the final results will be made available for you, but only after completion of the entire research report.

Confidentiality and Anonymity

Confidentiality is of utmost importance in all stages of this research. All data will be de-identified. This will include the removal of any names or other potentially identifying information you may mention in your interviews. No individual data will be described or released in any form, and only aggregate data will be presented in any reports based on these data. Your employer has given permission for employees to take part but will not be notified of your specific participation and will not be provided with the individual employee data.

Queries

Any queries or concerns regarding the research project can be addressed by contacting:

Researcher : Zhenan Feng
Phone : +64 9 373 7599 ext 88166
E-mail : zfen124@aucklanduni.ac.nz

Supervisor : Dr. Vicente Gonzalez
Phone : 09 3737599 ext 84106
E-mail : v.gonzalez@auckland.ac.nz

Co-Supervisor : Prof. Robert Amor
Phone : 09 3737599 ext 83068
E-mail : trebor@cs.auckland.ac.nz

Co-Supervisor : Assoc. Prof. Carol Mutch
Phone : 09 3737999 ext 48257
E-mail : c.mutch@auckland.ac.nz

For any queries regarding ethical concerns please contact: The Chair, University of Auckland Human Participants Ethics Committee Phone: 09 373 7599 ext. 83711. Postal Address: The University of Auckland, Office of the Vice-Chancellor, Private Bag 92019, Auckland 1142. E-Mail: ro-ethics@auckland.ac.nz

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON ________

13 /April/2016 FOR (3) YEARS REFERENCE NUMBER 016763

Appendix C: Consent Form

Department of Civil and Environmental Engineering
Engineering Building
20 Symonds Street
Auckland 1142
New Zealand
Telephone 64 9 373 7599 ext 88166
Facsimile 64 9 373 7462

The University of Auckland
Private Bag 92019
Auckland 1142
New Zealand

CONSENT FORM (For Ormiston Junior College Deputy Principal)

THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Research Project Title	Building Quake and People – A Serious Game Platform for Informing Life-Saving Strategies (Ref. 016763)
Researchers	Zhenan Feng, Civil and Environmental Engineering Department
Supervisor	Dr Vicente Gonzalez, Civil and Environmental Engineering Department
Co-Supervisor	Prof Robert Amor, Computer Science Department Assoc. Prof Carol Mutch, Education and Social Work Department

- We have read the Participant Information Sheet (PIS), we have understood the nature of this research and consent to the school's participation. We understand that our employees and students' participation is voluntary, and they may choose not to participate, even though the school has agreed to participate. Our employees and students have had the opportunities to ask questions and have had them answered to their satisfaction.
- We understand that our employees and students are free to withdraw their participation at any time, without any explanation.
- We understand that our employees and students will be observed during the use of the serious game and video recorded and they have the right to

withdraw the game and turn off the video recorder at any time during the serious game session, without the need to provide any reason.

- We understand that our employees and students will be interviewed after the use of the serious game and audio recorded and they have the right to withdraw the interview and turn off the recorder at any time during the interview session, without the need to provide any reason.
- We understand that upon request, our employees and students will have the right to review and edit the interview transcripts from an audio recording, to comply with Ormiston Junior College confidentiality requirement and will respond to it within two weeks from the receipt of the document.
- We understand that the data our employees and students provided will be stored securely within the university premises and only the researcher and supervisor may gain access to it.
- We understand that the data will be kept for six years, after which it will be destroyed.
- We understand that our employees and students should not reveal anything that is commercially sensitive to other third parties.
- We understand that our employees and students are free to withdraw any data traceable to them up to one month after undertaking the interview, to ensure that the level of reported information complies with Ormiston Junior College confidentiality requirement.
- We give assurance that our employees and students' participation or nonparticipation will not affect their relationship with the school or career/study status in any manner.
- We understand that our employees and students are entitled to request a copy of the final report.

Name : ________________________

Signature : ________________________ *Date :*______________________________

Please include your email address in the following space, if you would like to receive a copy of the final report. ______________________________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON _______

13 /April/2016 FOR (3) YEARS REFERENCE NUMBER 016763

Department of Civil and Environmental Engineering
Engineering Building
20 Symonds Street
Auckland 1142
New Zealand
Telephone 64 9 373 7599 ext 88166
Facsimile 64 9 373 7462

The University of Auckland
Private Bag 92019
Auckland 1142
New Zealand

CONSENT FORM (For Class Teacher)
THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Research Project Title	Building Quake and People – A Serious Game Platform for Informing Life-Saving Strategies (Ref. 016763)
Researchers	Zhenan Feng, Civil and Environmental Engineering Department
Supervisor	Dr Vicente Gonzalez, Civil and Environmental Engineering Department
Co-Supervisor	Prof Robert Amor, Computer Science Department Assoc. Prof Carol Mutch, Education and Social Work Department

- I have read the Participant Information Sheet (PIS), I have understood the nature of this research and consent to my students' participation. I understand that my students' participation is voluntary, and they may choose not to participate, even though the school has agreed to participate. My students have had the opportunities to ask questions and have had them answered to their satisfaction.
- I understand that my students are free to withdraw their participation at any time, without any explanation.
- I understand that my students will be observed during the use of the serious game and video recorded and they have the right to withdraw the game and

turn off the video recorder at any time during the serious game session, without the need to provide any reason.

- I understand that my students will be interviewed after the use of the serious game and audio recorded and they have the right to withdraw the interview and turn off the recorder at any time during the interview session, without the need to provide any reason.
- I understand that upon request, my students will have the right to review and edit the interview transcripts from an audio recording, to comply with Ormiston Junior College confidentiality requirement and will respond to it within two weeks from the receipt of the document.
- I understand that the data my students provided will be stored securely within the university premises and only the researcher and supervisor may gain access to it.
- I understand that the data will be kept for six years, after which it will be destroyed.
- I understand that my students should not reveal anything that is commercially sensitive to other third parties.
- I understand that my students are free to withdraw any data traceable to them up to one month after undertaking the interview, to ensure that the level of reported information complies with Ormiston Junior College confidentiality requirement.
- I give assurance that my students' participation or nonparticipation will not affect their relationship with me or study status in any manner.
- I understand that my students are entitled to request a copy of the final report.

Name : ________________________

Signature : ________________________ *Date :*______________________________

Please include your email address in the following space, if you would like to receive a copy of the final report. ______________________________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON _______

13 /April/2016 FOR (3) YEARS REFERENCE NUMBER 016763

Department of Civil and Environmental Engineering
Engineering Building
20 Symonds Street
Auckland 1142
New Zealand
Telephone 64 9 373 7599 ext 88166
Facsimile 64 9 373 7462

The University of Auckland
Private Bag 92019
Auckland 1142
New Zealand

CONSENT FORM (For Guardians)

THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Research Project Title	Building Quake and People – A Serious Game Platform for Informing Life-Saving Strategies (Ref. 016763)
Researchers	Zhenan Feng, Civil and Environmental Engineering Department
Supervisor	Dr Vicente Gonzalez, Civil and Environmental Engineering Department
Co-Supervisor	Prof Robert Amor, Computer Science Department Assoc. Prof Carol Mutch, Education and Social Work Department

- I understand that my child agrees to voluntary take part in this research. I have read the Participant Information Sheet (PIS), I have understood the nature of this research and why my child has been selected. I have had the opportunity to ask questions and have had them answered to my satisfaction.
- I understand that my child is free to withdraw his/her participation at any time, without any explanation.
- I understand that my child will be observed during the use of the serious game and video recorded and he/she has the right to withdraw the game and turn off the video recorder at any time during the serious game session, without the need to provide any reason.

- I understand that my child will be interviewed after the use of the serious game and audio recorded and he/she has the right to withdraw the interview and turn off the recorder at any time during the interview session, without the need to provide any reason.
- I understand that upon request, my child will have the right to review and edit the interview transcripts from an audio recording, to comply with his/her organization's confidentiality requirement and will respond to it within two weeks from the receipt of the document.
- I understand that the data my child provided will be stored securely within the university premises and only the researcher and supervisor may gain access to it.
- I understand that the data will be kept for six years, after which it will be destroyed.
- I understand that my child should not reveal anything that is commercially sensitive to other third parties.
- I understand that my child is free to withdraw any data traceable to him/her up to one month after undertaking the interview, to ensure that the level of reported information complies with his/her organization's confidentiality requirement.
- I understand that as a guardian I have permitted to allow my child participating in this research.
- I understand that my child is entitled to request a copy of the final report.

Name : ________________________

Signature : ________________________ *Date :*______________________________

Please include your email address in the following space, if you would like to receive a copy of the final report. ______________________________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON _______

13 /April/2016 FOR (3) YEARS REFERENCE NUMBER 016763

Department of Civil and Environmental Engineering
Engineering Building
20 Symonds Street
Auckland 1142
New Zealand
Telephone 64 9 373 7599 ext 88166
Facsimile 64 9 373 7462

The University of Auckland
Private Bag 92019
Auckland 1142
New Zealand

CONSENT FORM (For Children)
THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Research Project Title	Building Quake and People – A Serious Game Platform for Informing Life-Saving Strategies (Ref. 016763)
Researchers	Zhenan Feng, Civil and Environmental Engineering Department
Supervisor	Dr Vicente Gonzalez, Civil and Environmental Engineering Department
Co-Supervisor	Prof Robert Amor, Computer Science Department Assoc. Prof Carol Mutch, Education and Social Work Department

- I agree to voluntary take part in this research. I have read the Participant Information Sheet (PIS). I have understood the nature of this research and why I have been selected. I have had the opportunity to ask questions and have had them answered to my satisfaction.
- I understand that I am free to withdraw my participation at any time, without any explanation.
- I understand that I will be observed during the research. My behaviour will be video recorded. I have the right to withdraw the research and turn off the recorder at any time without any reason.

- I understand that I will be interviewed after the research. My answer will be voice recorded. I have the right to withdraw the interview and turn off the recorder at any time without any reason.
- I understand that I have the right to review and edit the interview transcripts from audio recording upon request up to two weeks after undertaking the interview.
- I understand that the data I provide will be stored securely within the university premises and only the researcher and supervisor may gain access to it.
- I understand that the data will be kept for six years, after which it will be destroyed.
- I understand that I should not tell anything that is commercially sensitive to anyone else.
- I understand that I am free to withdraw any data traceable to me up to one month after undertaking the interview.
- I understand that my guardians and teachers have permitted to allow me participating in this research.
- I understand that I am entitled to request a copy of the final report.

Name : ________________________

Signature : ________________________ *Date :*______________________________

Please include your email address in the following space, if you would like to receive a copy of the final report. ______________________________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON _______

13 /April/2016 FOR (3) YEARS REFERENCE NUMBER 016763

Department of Civil and Environmental Engineering
Engineering Building
20 Symonds Street
Auckland 1142
New Zealand
Telephone 64 9 373 7599 ext 88166
Facsimile 64 9 373 7462

The University of Auckland
Private Bag 92019
Auckland 1142
New Zealand

CONSENT FORM
THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Research Project Title	Building Quake and People – A Serious Game Platform for Informing Life-Saving Strategies (Ref. 016763)
Researchers	Zhenan Feng, Civil and Environmental Engineering Department
Supervisor	Dr Vicente Gonzalez, Civil and Environmental Engineering Department
Co-Supervisor	Prof Robert Amor, Computer Science Department Assoc. Prof Carol Mutch, Education and Social Work Department

- I agree to voluntary take part in this research. I have read the Participant Information Sheet (PIS), I have understood the nature of this research and why I have been selected. I have had the opportunity to ask questions and have had them answered to my satisfaction.
- I understand that I am free to withdraw my participation at any time, without any explanation.
- I understand that I will be observed during the use of the serious game and video recorded and I have the right to withdraw the game and turn off the video recorder at any time during the serious game session, without the need to provide any reason.

- I understand that I will answer questionnaires before and after the use of the serious game and I have the right to withdraw the questionnaire without the need to provide any reason.
- I understand that the data I provide will be stored securely within the university premises and only the researcher and supervisor may gain access to it.
- I understand that the data will be kept for six years, after which it will be destroyed.
- I understand that I am free to withdraw any data traceable to me up to one month after undertaking the experiment, to ensure that the level of reported information complies with my organization's confidentiality requirement.
- I understand that I am entitled to request a copy of the final report.

Name : ________________________

Signature : ________________________ *Date :*_____________________________

Please include your email address in the following space, if you would like to receive a copy of the final report. ______________________________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON _______

13 /April/2016 FOR (3) YEARS REFERENCE NUMBER 016763