



Exploring spiral narratives with immediate feedback in immersive virtual reality serious games for earthquake emergency training

Zhenan Feng^{1,2}  · Vicente A. González¹ · Carol Mutch³ · Robert Amor⁴ · Guillermo Cabrera-Guerrero⁵

Received: 13 January 2021 / Revised: 31 May 2021 / Accepted: 30 May 2022
© The Author(s) 2022

Abstract

Various attempts and approaches have been made to teach individuals about the knowledge of best practice for earthquake emergencies. Among them, Immersive Virtual Reality Serious Games (IVR SGs) have been suggested as an effective tool for emergency training. The notion of IVR SGs is consistent with the concept of problem-based gaming (PBG), where trainees interact with games in a loop of forming a playing strategy, applying the strategy, observing consequences, and making reflection. PBG triggers reflection-on-action, enabling trainees to reform perceptions and establish knowledge after making a response to a scenario. However, in the literature of PBG, little effort has been made for trainees to reflect while they are making a response (i.e., reflection-in-action) in a scenario. In addition, trainees do not have the possibility to adjust their responses and reshape their behaviors according to their reflection-

✉ Zhenan Feng
z.fengl@massey.ac.nz

Vicente A. González
v.gonzalez@auckland.ac.nz

Carol Mutch
c.mutch@auckland.ac.nz

Robert Amor
trebor@cs.auckland.ac.nz

Guillermo Cabrera-Guerrero
guillermo.cabrera@pucv.cl

¹ Department of Civil and Environmental Engineering, The University of Auckland, Auckland, New Zealand

² School of Built Environment, Massey University, Auckland, New Zealand

³ Department of Critical Studies in Education, The University of Auckland, Auckland, New Zealand

⁴ Department of Computer Science, The University of Auckland, Auckland, New Zealand

⁵ Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

in-action. In order to overcome these limitations, this study proposes a game mechanism, which integrates spiral narratives with immediate feedback, to underpin reflection-in-action and reflective redo in PBG. An IVR SG training system suited to earthquake emergency training was developed, incorporating the proposed game mechanism. A controlled experiment with 99 university students and staff was conducted. Participants were divided into three groups, with three interventions tested: a spiral narrated IVR SG, a linear narrated IVR SG, and a leaflet. Both narrated IVR SGs were effective in terms of immediate knowledge gain and self-efficacy improvement. However, challenges and opportunities for future research have been suggested.

Keywords Immersive Virtual Reality · Serious Games · Earthquake training · Problem-based gaming · Narrative

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 [62]. Suitable and timely behavioral responses are essential to save lives and reduce injuries during earthquake emergencies [5]. In recent years, Immersive Virtual Reality Serious Games (IVR SGs) have increasingly gained attention for education and training [27]. IVR offers credible virtual environments that engage users through a high level of immersion [21, 56]. In the literature, IVR SGs have been reported to deliver quality training for emergencies [19].

Currently, most of the IVR SGs for emergency training have applied reflection-on-action, in which case trainees reflect on their actions after completing an activity. Little attention has been put on reflection-in-action. Reflection-in-action refers to the reflection that takes place while individuals are carrying out an activity [30]. Individuals need real-time feedback to make sense of confronting situations, adapt themselves, adjust behaviors, and take actions accordingly [30]. Reflection-in-action shapes the output of individuals, in most cases, modifying their behavioral responses to current situations [30]. Therefore, the facilitation of reflection-in-action is important to train individuals about behavioral responses to specific scenarios.

This study aims to investigate the integration of reflection-in-action and reflective redo with 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. How effective is an IVR SG with reflection-in-action and reflective redo in terms of training?

The following research objectives were established in order to address the research questions:

1. To understand IVR SGs for emergency training, storylines and narratives in GBL, and reflection and redo in GBL.
2. To develop an IVR SG for earthquake emergency training with reflection-in-action and reflective redo, using spiral narratives and immediate feedback.
3. To validate the effectiveness of training for the proposed IVR SG with reflection-in-action and reflective redo.

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.

2 State of the art

2.1 IVR SGs and problem-based gaming

IVR SGs enables training to take place in credible virtual environments. As such, IVR SGs have become popular for emergency training [19], covering a range of emergency situations, such as fire [57], aircraft evacuation [10, 13], and earthquakes [38]. Smith and Ericson [57] have motivated children to learn about fire evacuation skills using a projection-based IVR SG. Trainees have acknowledged the value of IVR SGs due to the fact they provide highly engaging, interactive training environments and experiences. Burigat and Chittaro [10] put trainees in an IVR SG to learn and improve their spatial knowledge associated with aircrafts. Results indicated that the IVR SG outperformed the traditional training approach (i.e. safety cards) in terms of teaching spatial knowledge. Through the IVR SG training environment, trainees had the chance to walk in an aircraft and get familiar with its spatial layout. Trainees paid more attention to learning tasks in the IVR SG than using safety cards. In another aircraft emergency training, trainees learned the appropriate behavioral responses to aircraft emergency landing and evacuation through an IVR SG [13]. Results indicated that trainees who were trained in the IVR SG were able to retain more learned safety knowledge than those trained with safety cards. The possible reason is the high-level emotional arousal induced by realistic simulations (especially during hazardous events) within an IVR SG environment. Regarding earthquake training, Li et al. [38] proposed an IVR SG to teach self-protection skills in an earthquake. Results revealed that the IVR SG performed better than other traditional approaches, including videos and safety manuals. Li et al. [38] suggested that realistic experience and immersive practice could overcome the gap between theory and practice existing in traditional training approaches. In sum, IVR SGs seem to be a promising and effective training tool for emergency situations.

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 [36]. Figure 1 illustrates the concept of PBG. According to Kiili [36], 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 [6]. Reflection is a vital part of PBG as it constructs cognition and synthesizes knowledge [36]. 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 [36].

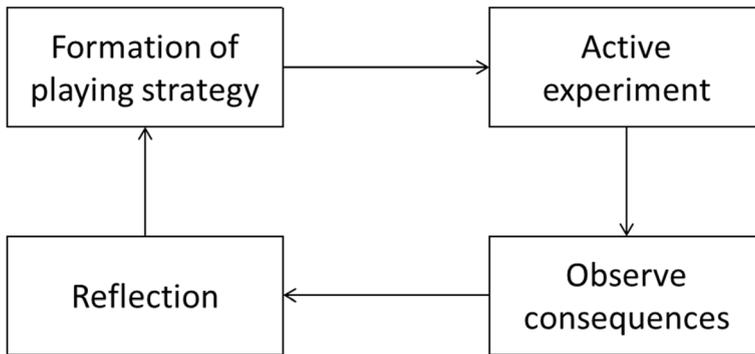


Fig. 1 Problem-based gaming, derived from [36]

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.

2.2 Storylines and narratives

Fundamentally, an IVR SG incorporates the concept of game-based learning (GBL), which covers three essential aspects: learning, play, and storylines [29]. A storyline is a game element that presents game content with structured events [32]. In GBL, learners engage with storylines to play and learn. Knowledge is conveyed in context through storylines within game environments [59]. The use of storylines is beneficial to improving the understanding and recall of presented materials [7]. In addition, storylines promote the immersion and motivation of learners [46]. Appealing storylines encourage learners to engage in gameplay, leading to enhanced learning effects [31]. As such, storylines set out a baseline in GBL, with a strong influence on other gaming characteristics, such as interactivity, challenges, and feedback [2].

Storylines consist of a set of scenarios. The way to serve these scenarios to learners is referred to as narratives (i.e., storytelling methods) [12]. Narratives are essential to communicate learning content to learners [39]. In principle, there are two types of narratives for GBL: an implicit narrative and an explicit narrative [25]. In the case of implicit narratives, the storylines are not structurally outlined. The presentation of storylines depends on the exploration of learners in game environments. Regarding explicit narratives, storylines are clearly presented to learners. Storylines take the lead in learner activities. Implicit narratives may lead to an increased recall of spatial knowledge, whereas explicit narratives may perform better on transferring factual knowledge [25]. 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 [22]. 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, 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.

2.3 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 [42]. There are two categories of reflection: reflection-in-action and reflection-on-action [50]. Reflection-in-action means the reflection occurs while individuals are responding to a situation [30]. With the effort of making sense of confronting situations, individuals reflect on their understandings and make responses accordingly [30]. Reflection-in-action reshapes the perceptions of individuals and guides individuals’ follow-on actions [30]. In terms of reflection-on-action, it is a retrospective process that 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 [30]. 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 [24]. 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 [33]. When learners interact with game 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 [50]. 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 on 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 [53]. 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 [53]. In the literature of GBL, reflective redo has been studied for After Action Review (AAR) [53]. AAR is a retrospective process that provides feedback after training exercises, similar to reflection-on-action [3]. Scoresby and Shelton [53] 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.

3 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 practices 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.

3.1 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 1) [45]. 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 [59]. An action-driven approach was applied to drive the storyline. The storyline was progressed once trainees had taken actions in scenarios [22]. In our case, trainees had to select options to solve problems confronted in the IVR SG as the storyline goes on.

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 [17]. In the case of IVR SGs, with the integration of an action-driven approach, storylines progress from one scenario

Table 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.

to the next after trainees take actions to solve a problem in a scenario (see Fig. 2a). 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 Fig. 2b). 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.

3.2 Immediate feedback

In order to foster reflection, immediate feedback is given to trainees after a specific response is undertaken [36]. 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 Fig. 3b, c). Audio feedback stands for a sound effect that 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 Fig. 3d). 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 [9]. 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

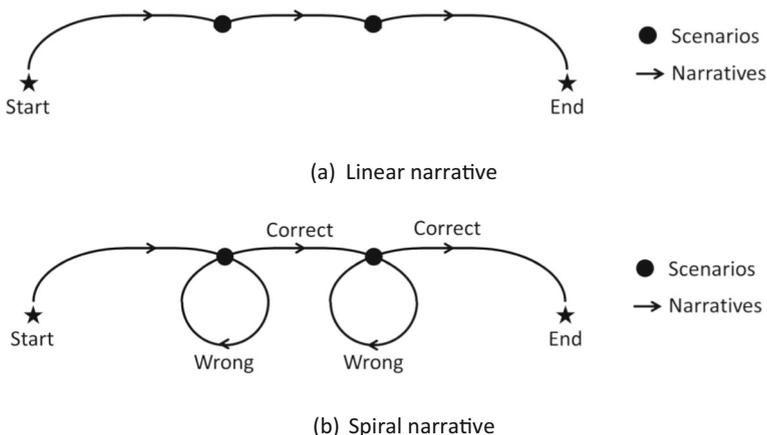


Fig. 2 Linear narrative and spiral narrative. (a) Linear narrative. (b) Spiral narrative

applied for positive feedback, with all 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 [9, 11].

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 [49]. 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 [37]. 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 [13]. 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

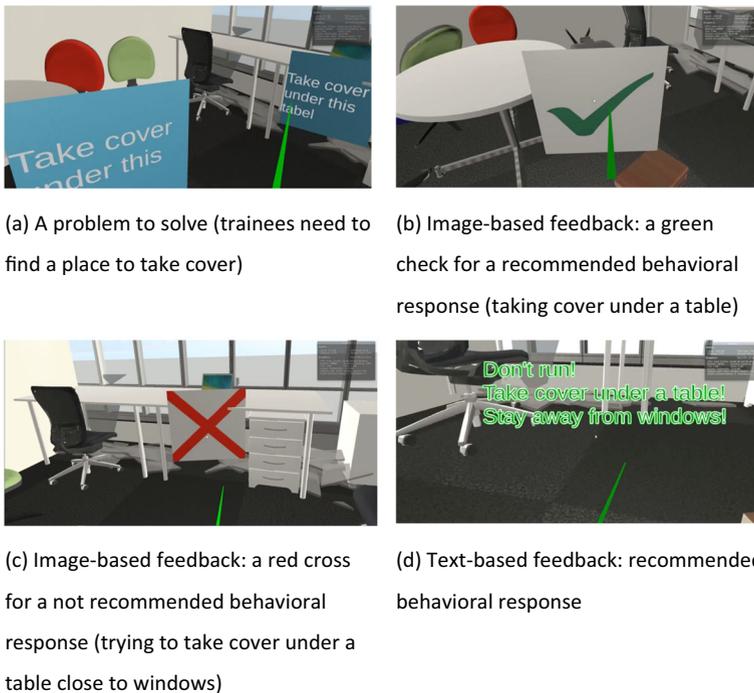


Fig. 3 Immediate feedback deployed in the IVR SG training system. (a) A problem to solve (trainees need to find a place to take cover). (b) Image-based feedback: a green check for a recommended behavioral response (taking cover under a table). (c) Image-based feedback: a red cross for a not recommended behavioral response (trying to take cover under a table close to windows). (d) Text-based feedback: recommended behavioral response

time was available for trainees when they received immediate feedback [10]. 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 [13]. 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.

3.3 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. [41] 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 [20, 41]. A basic building model defining the envelope and layout of the selected built environment was developed using Autodesk Revit (see Fig. 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) [55].

A qualitative approach was applied to model an earthquake and provide trainees with the sense of being in an actual earthquake [41]. 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 [23]. In our case, we referred to the New Zealand Modified Mercalli Scale (MMI) to simulate an earthquake and its damage [28]. We selected the description of

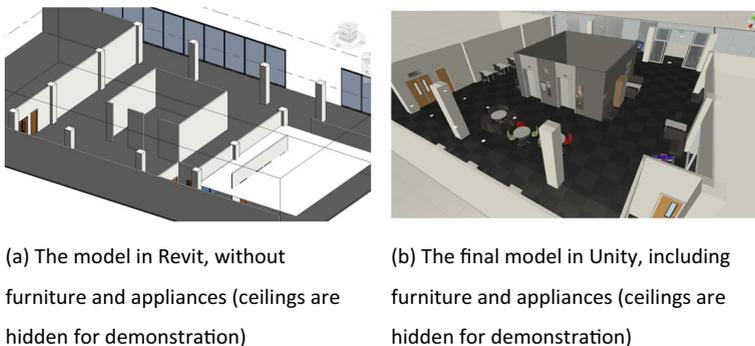


Fig. 4 The models in Revit and Unity. (a) The model in Revit, without furniture and appliances (ceilings are hidden for demonstration). (b) The final model in Unity, including furniture and appliances (ceilings are hidden for demonstration)

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” [28]. 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. Figure 5 demonstrates the virtual environment and earthquake damage.

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.

Latency is an important factor in IVR systems, which could lead to a loss of performance and cybersickness [60]. In order to address latency, we applied a holistic approach, including software design and hardware configuration. Firstly, we used low-polygon models to decrease the rendering demand of the IVR environment, as described in this section. Secondly, we used Occlusion Culling in Unity to disable rendering of objects that are currently not visible by cameras (i.e., participants). Thirdly, we used a high-spec computer to run the IVR SG, as described in this section. Fourthly, we limited the movement of participants on a predefined route, using a waypoint navigation system [47]. Participants clicked a button on their remote controller, and the camera started to move from one waypoint to the next, following the route. In addition, participants remained seated during the entire IVR experience. They only needed to turn their bodies via a swivel chair to face their moving direction. Therefore, participants were not always paying attention to making movement and interaction in the IVR environment. Taken together, the first three approaches were aimed to reduce the latency itself by increasing frames per second (FPS). The last approach tried to reduce the impact of latency on participants by involving less interaction.

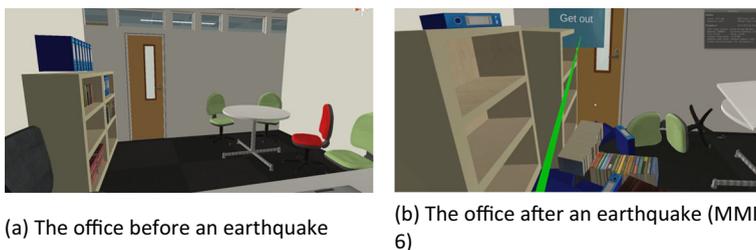


Fig. 5 Virtual environment and earthquake damage. **(a)** The office before an earthquake. **(b)** The office after an earthquake (MMI 6)

4 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 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.

4.1 Materials

The Spiral narrative and Linear narrative groups were treated with the IVR SG training system, using the software and hardware described in Section 3.3. 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 1.

4.2 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 2). No significant differences were revealed between groups based on Kruskal-Wallis tests (earthquake drills, $p = 0.593$; IVR, $p = 0.523$).

4.3 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 [23]. Following a pretest-posttest research design, a questionnaire (see Sections 4.3.1 and 4.3.2) 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 4.3.3). We also measured the

Table 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%

perceptions of trainees towards the deployed narrative and feedback mechanisms in the IVR SG training system (see Section 4.3.4), the levels of cybersickness (see Section 4.3.5), and the realism of the virtual environments (see Section 4.3.6).

4.3.1 Safety knowledge

Trainees are expected to be able to deal with similar real-life scenarios after training [38]. 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 1). In order to measure the acquisition of safety knowledge, a true-false knowledge test was established containing ten questions (see Table 3). The true-false knowledge test was aimed to evaluate whether participants understood what appropriate behavioral responses in earthquake emergencies are. The true-false type of questions has been applied in the literature for knowledge tests [35]. 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 3 illustrates the statements and the correct answers to the knowledge test.

4.3.2 Self-efficacy

Self-efficacy means the beliefs people hold in their competency to solve problems and overcome difficulties [4]. High-level self-efficacy may result in a change in behavior, which leads to the improvement of performance when dealing with problems and difficulties [58]. The General Self-Efficacy Scale has been suggested to measure self-efficacy in the literature, providing a list of statements [52]. 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."

Table 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

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 [52]. In our case, the possible total score ranged from -18 to 18 .

4.3.3 Attention

For IVR groups, we deployed a self-reported questionnaire to assess to what extent the training approaches attracted trainees' attention. The questionnaire included three statements, following the measurements applied by Burigat and Chittaro [10]:

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 [10].

4.3.4 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 [14], we developed a questionnaire for IVR groups, 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 [14].

4.3.5 Cybersickness

As discussed in Section 3.3, we applied a holistic approach to address latency issues. In order to measure cybersickness, for IVR groups, we used the following statement for trainees to rate:

1. "The VR experience made me dizzy".

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.

4.3.6 Realism

Another aspect of user experience measured in this study was the realism of virtual environments. For IVR groups, we used the following statements for trainees to rate:

1. “The building environment was realistic”;
2. “The VR experience was realistic”.

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.

4.4 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 that 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 of 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.

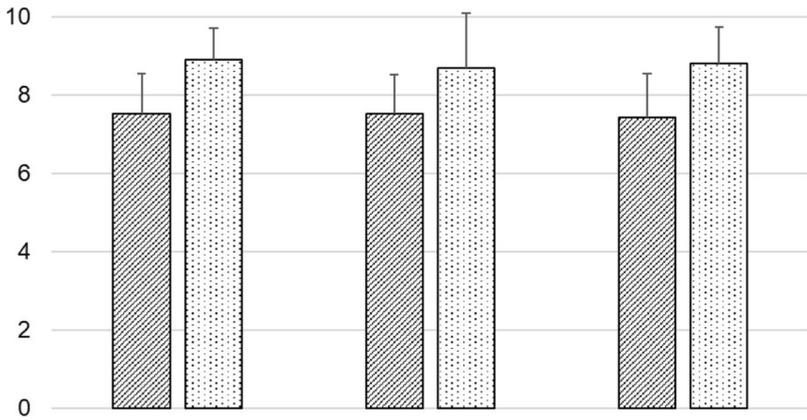
Regarding the leaflet group, trainees were trained through reading a leaflet. Trainees were instructed to study the leaflet carefully till they fully understood the content, no matter how long it took.

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.

5 Results

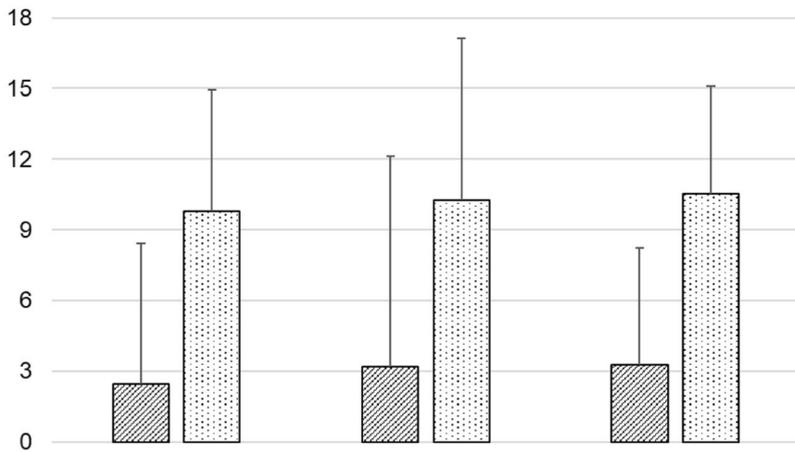
The results of knowledge scores and self-efficacy scores are reported in Fig. 6. Wilcoxon Signed Ranks Tests confirmed that the knowledge and self-efficacy were improved significantly after training within each group. ANCOVA was adopted to analyse the pre-test and post-test scores between groups. The 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 4.

Knowledge



	Linear narrative	Spiral narrative	Leaflet
Pre-test	M = 7.52, SD = 1.03	M = 7.52, SD = 1.00	M = 7.42, SD = 1.12
Post-test	M = 8.91, SD = 0.80	M = 8.70, SD = 1.40	M = 8.82, SD = 0.92
Significance	p < 0.001	p < 0.001	p < 0.001

Self-efficacy



	Linear narrative	Spiral narrative	Leaflet
Pre-test	M = 2.45, SD = 5.96	M = 3.21, SD = 8.92	M = 3.27, SD = 4.94
Post-test	M = 9.79, SD = 5.16	M = 10.24, SD = 6.84	M = 10.55, SD = 4.54
Significance	p < 0.001	p < 0.001	p < 0.001

Fig. 6 Pre-test and post-test results of knowledge and self-efficacy. Error bars indicate the standard deviation of the mean

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

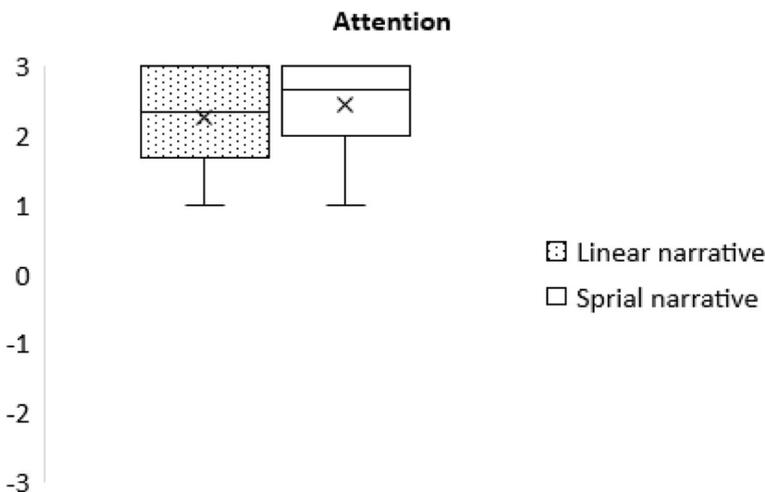
Pairwise comparisons	Knowledge acquisition	Self-efficacy improvement
Leaflet vs. Linear narrative	No significant difference	No significant difference
Leaflet vs. Spiral narrative	No significant difference	No significant difference
Linear narrative vs. Spiral narrative	No significant difference	No significant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

The results of attention are reported in box plots (Linear narrative: $M = 2.26$, $SD = 0.70$; Spiral narrative: $M = 2.44$, $SD = 0.63$), as shown in Fig. 7. Cronbach's alphas were calculated to assess the internal consistency of multiple statements (Linear narrative: 0.868; Spiral narrative: 0.778), suggesting that the three asked statements were closely related to measuring attention [15]. Lastly, Kruskal-Wallis tests confirmed no significant difference between groups ($p = 0.282$). In general, the trainees in both IVR groups acknowledged that they kept engaged with the IVR SG.

The results of ease of training are reported in box plots (Linear narrative: $M = 2.28$, $SD = 0.710$; Spiral narrative: $M = 2.33$, $SD = 0.741$), as shown in Fig. 8. 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 measuring ease of training. Lastly, Kruskal-Wallis tests confirmed no significant difference between groups ($p = 0.633$). In general, the trainees in both IVR groups acknowledged that the IVR SG was easy for them to get training.

No trainees quit the experiment because of cybersickness. The ratings of cybersickness are reported in box plots (Linear narrative: $M = -0.82$, $SD = 0.2.02$; Spiral narrative: $M = -0.10$, $SD = 1.94$), as shown in Fig. 9. Kruskal-Wallis tests confirmed no significant difference between groups ($p = 0.140$). In general, the trainees in both IVR groups acknowledged that the IVR SG did not cause serious cybersickness issues.

The ratings of realism are reported in box plots (Linear narrative: $M = 2.14$, $SD = 0.753$; Spiral narrative: $M = 2.21$, $SD = 0.977$), as shown in Fig. 10. Cronbach's alphas were

**Fig. 7** The attention levels reported by trainees

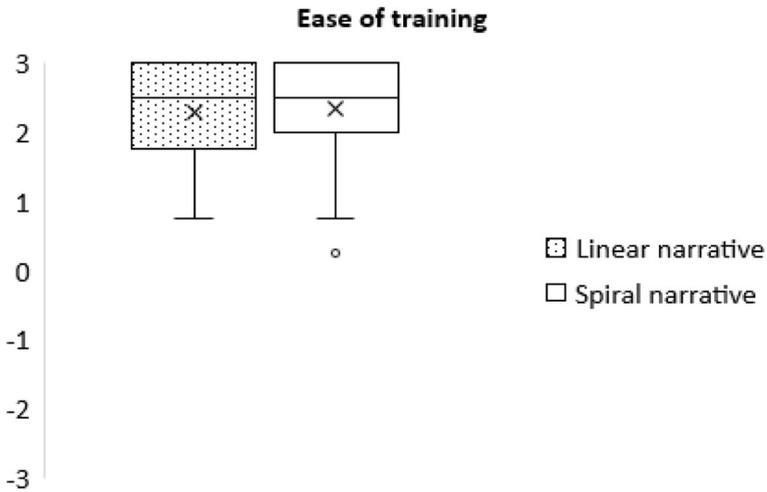


Fig. 8 The ease of training reported by trainees

calculated to assess the internal consistency of multiple statements (Linear narrative: 0.900; Spiral narrative: 0.934), suggesting that the two asked statements were closely related to measuring realism. Lastly, Kruskal-Wallis tests confirmed no significant difference between groups ($p = 0.336$). In general, the trainees in both IVR groups acknowledged that the IVR SG provided high levels of realism in terms of virtual environments.

6 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.

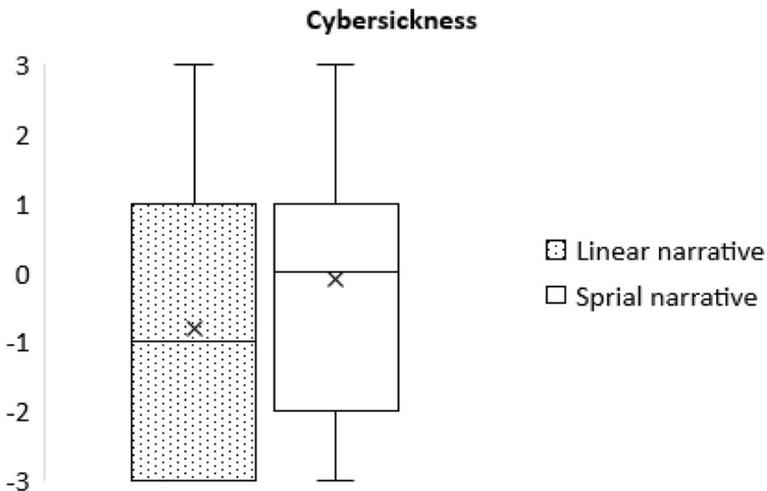


Fig. 9 The cybersickness reported by trainees

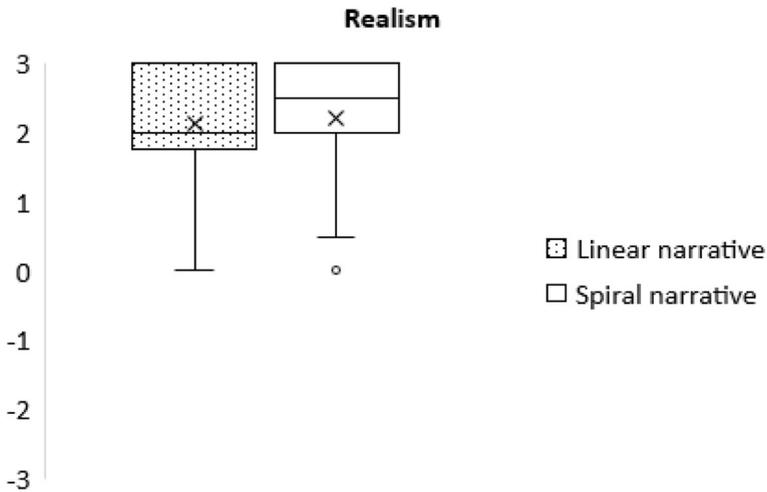


Fig. 10 The realism reported by trainees

6.1 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 instructional 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 [10, 23].

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.

6.2 IVR SG training vs. leaflet

The three tested approaches were all effective in enhancing earthquake preparedness, manifesting in significant improvement in 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 [13, 57], 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 [13] 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 [26, 34, 54]. Traditional approaches, such as safety cards or leaflets, are incapable of arousing intense emotion [13].

6.3 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 [40]. The same effect might occur in 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 [1]. Flow is a state in which trainees are highly engaged in activities and functions to manage activities [16, 43]. Flow plays an important role in GBL, by keeping trainees concentrating on learning [1]. 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 [53]. Metacognition is referred to as "the deliberate conscious control of one's own cognitive actions" [8], in other words, the knowledge that people hold about their thoughts. Scoresby and Shelton [53] 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 [61]. 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 practising in expansive contexts are beneficial for trainees to transferring knowledge and applying skills to new settings in the future [18, 44], which is essential for earthquake safety as it is spatially unpredictable in times of need.

Another important factor to be investigated in future is the physiological and psychological aspects of participants undergoing training in IVR environments. Adverse effects may occur to participants as IVR may lead to disorientation, nausea, anxiety, flashbacks, and post-traumatic stress disorder (PTSD) [48, 51]. The potential issues raised by an IVR training system must be investigated, and the responding procedures must be in place before rolling out to a larger audience.

7 Conclusions

IVR SGs have been studied for earthquake and emergency training in the literature. However, these IVR studies mainly focused on the prototyping aspect, with little attention to narratives, reflection, and their impacts on training. The present study contributes to the body of knowledge by incorporating reflection-in-action and reflective redo with IVR SGs for earthquake training. The proposed IVR SG features immediate feedback and spiral narratives. With the manipulation of narratives, trainees are allowed to repeat a scenario with immediate feedback to induce reflection-in-action, enabling reflective redo. A controlled experiment was conducted to validate the proposed game mechanisms. An IVR SG with linear narratives and an IVR SG with spiral narratives were compared, with a leaflet approach being a control group. Immediate knowledge gain, self-efficacy improvement, attention, ease of training, cybersickness, and realism of the IVR SG were measured. Results support that the IVR SG training system is well aligned to PBG. Trainees 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 immediately after training. IVR SGs are likely to have a long-term impact on memory [13]. 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 s 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. 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.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data availability Yes.

Code availability No. There are commercial concerns.

Declarations

Conflict of interest No conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or

exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Admiraal W, Huizenga J, Akkerman S et al (2011) The concept of flow in collaborative game-based learning. *Comput Hum Behav* 27(3):1185–1194
2. Asgari M, Kaufman D (2004) Relationships among computer games, fantasy, and learning. Simon Fraser University
3. Baird L, Holland P, Deacon S (1999) Learning from action: Imbedding more learning into the performance fast enough to make a difference. *Organ Dyn* 27(4):19–32
4. Bandura A (2010) Self-efficacy. Anonymous The Corsini encyclopedia of psychology. Wiley Online Library, pp 1–3
5. Bernardini G, D’Orazio M, Quagliarini E (2016) Towards a “behavioural design” approach for seismic risk reduction strategies of buildings and their environment. *Saf Sci* 86:273–294
6. Boud D, Keogh R, Walker D (2013) Reflection: Turning experience into learning. Routledge, Abingdon
7. Bransford JD, Sherwood RD, Hasselbring TS et al (2012) Anchored instruction: Why we need it and how technology can help. Anonymous Cognition, education, and multimedia. Routledge, Abingdon, pp 129–156
8. Brown AL (1980) Metacognitive development and reading. In: Spiro RJ, Bruce BB, Brewer WF (eds.), *Theoretical Issues in Reading Comprehension*. Routledge, London, pp 453–481. <https://www.taylorfrancis.com/books/edit/10.4324/9781315107493/theoretical-issues-readingcomprehension-rand-spiro-bertram-bruce-william-brewer>
9. Brown AS, Schilling HE, Hockensmith ML (1999) The negative suggestion effect: Pondering incorrect alternatives may be hazardous to your knowledge. *J Educ Psychol* 91(4):756
10. Burigat S, Chittaro L (2016) Passive and active navigation of virtual environments vs. traditional printed evacuation maps: A comparative evaluation in the aviation domain. *Int J Hum - Comput Stud* 87:92–105. <https://doi.org/10.1016/j.ijhcs.2015.11.004>
11. Butler AC, Roediger HL (2008) Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem Cognit* 36(3):604–616
12. Chatman SB (1980) *Story and discourse: Narrative structure in fiction and film*. Cornell University Press, Ithaca
13. Chittaro L, Buttussi F (2015) Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE Trans Vis Comput Graph* 21(4):529–538. <https://doi.org/10.1109/TVCG.2015.2391853>
14. Chittaro L, Sioni R (2015) Serious games for emergency preparedness: Evaluation of an interactive vs. a non-interactive simulation of a terror attack. *Comput Hum Behav* 50:508–519. <https://doi.org/10.1016/j.chb.2015.03.074>
15. Cohen BH (2013) *Explaining psychological statistics*, 4th edn. Wiley, Hoboken 2013. ©2014
16. Csikszentmihalyi M (1990) *Flow: The psychology of optimal experience*. Harper & Row, New York
17. Daniel C, Rod M (2020) *A dictionary of media and communication*, 3rd edn. Oxford University Press, Oxford
18. Engle RA, Lam DP, Meyer XS et al (2012) How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychol* 47(3):215–231
19. Feng Z, González VA, Amor R et al (2018) Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Comput Educ* 127:252–266. <https://doi.org/10.1016/j.compedu.2018.09.002>
20. Feng Z, González VA, Ma L et al (2018) Rapid 3D reconstruction of indoor environments to generate virtual reality serious games scenarios. In: Anonymous 18th International Conference on Construction Applications of Virtual Reality. University of Auckland, Auckland
21. Feng Z, González VA, Trotter M et al (2020) How people make decisions during earthquakes and post-earthquake evacuation: Using verbal protocol analysis in immersive virtual reality. *Saf Sci* 129:104837. <https://doi.org/10.1016/j.ssci.2020.104837>
22. Feng Z, González VA, Mutch C et al (2020) Towards a customizable immersive virtual reality serious game for earthquake emergency training. *Adv Eng Inform* 46:101134

23. Feng Z, González VA, Amor R et al (2020) An immersive virtual reality serious game to enhance earthquake behavioral responses and post-earthquake evacuation preparedness in buildings. *Adv Eng Inform* 45:101118. <https://doi.org/10.1016/j.aei.2020.101118>
24. Feng Z, González VA, Mutch C et al (2021) Instructional mechanisms in immersive virtual reality serious games: Earthquake emergency training for children. *J Comput Assist Learn* 2(37):542–556. <https://doi.org/10.1111/jcal.12507>
25. Ferguson C, van den Broek EL, van Oostendorp H (2020) On the role of interaction mode and story structure in virtual reality serious games. *Comput Educ* 143:103671
26. Finn B, Roediger HL III (2011) Enhancing retention through reconsolidation: Negative emotional arousal following retrieval enhances later recall. *Psychol Sci* 22(6):781–786
27. Freina L, Ott M (2015) A literature review on immersive virtual reality in education: State of the art and perspectives. In: Anonymous The International Scientific Conference eLearning and Software for Education
28. GeoNet (2019) Earthquake intensity. Available via <https://www.geonet.org.nz/earthquake/intensity>. Accessed 13/11/ 2019
29. Göbel S, Wendel V, Ritter C et al (2010) Personalized, adaptive digital educational games using narrative game-based learning objects. In: Anonymous 5th International Conference on E-learning and Games, Edutainment 2010. Springer, Berlin, p 438
30. Greenwood J (1993) Reflective practice: a critique of the work of Argyris and Schön. *J Adv Nurs* 18(8): 1183–1187
31. Habgood MJ, Ainsworth SE, Benford S (2005) Endogenous fantasy and learning in digital games. *Simul Gaming* 36(4):483–498
32. Ippa N, Borst T (2012) Story and simulations for serious games: Tales from the trenches. CRC Press, Boca Raton
33. Johnson CI, Priest HA (2014) The feedback principle in multimedia learning. Cambridge University Press, Cambridge
34. Kensinger EA (2009) Remembering the details: Effects of emotion. *Emot Rev* 1(2):99–113
35. Khan S, Davies DA, Gupta JK (2001) Formative self-assessment using multiple true-false questions on the Internet: feedback according to confidence about correct knowledge. *Med Teach* 23(2):158–163
36. Kiili K (2007) Foundation for problem-based gaming. *Br J Edu Technol* 38(3):394–404
37. Krathwohl DR, Anderson LW (2009) A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives. Longman, Harlow
38. Li C, Liang W, Quigley C et al (2017) Earthquake safety training through virtual drills. *TVCG* 23(4):1275–1284. <https://doi.org/10.1109/TVCG.2017.2656958>
39. Lim T, Louchart S, Suttie N et al (2014) Narrative serious game mechanics (NSGM)–insights into the narrative-pedagogical mechanism. In: Anonymous International conference on serious games. Springer, Berlin, p 23
40. 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. *Adv Eng Inform* 39:53–61
41. Lovreglio R, González VA, Feng Z et al (2018) Prototyping virtual reality serious games for building earthquake preparedness: The Auckland City Hospital case study. *Adv Eng Inform* 38:670–682. <https://doi.org/10.1016/j.aei.2018.08.018>
42. Moon JA (2013) Reflection in learning and professional development: Theory and practice. Routledge, Abingdon
43. Nakamura J, Csikszentmihalyi M (2014) The concept of flow. Anonymous flow and the foundations of positive psychology. Springer, New York, pp 239–263
44. National Research Council (2000) How people learn: Brain, mind, experience, and school: Expanded edition. National Academies Press, Washington, DC
45. New Zealand Ministry of Civil Defence & Emergency Management (2015) Working from the same page consistent messages for CDEM - Part B. Hazard specific information - Earthquake. Available via <https://www.civildefence.govt.nz/cdem-sector/consistent-messages-for-cdem/>. Accessed 13 Nov 2019
46. Padilla-Zea N, Gutiérrez FL, López-Arcos JR et al (2014) Modeling storytelling to be used in educational video games. *Comput Hum Behav* 31:461–474
47. Ragavan SV, Ponnambalam SG, Sumero C (2011) Waypoint-based path planner for mobile robot navigation using PSO and GA-AIS. In: Anonymous Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE. IEEE, p 756
48. Regan C (1995) An investigation into nausea and other side-effects of head-coupled immersive virtual reality. *Virtual Reality* 1(1):17–31
49. Romano DM, Brna P (2001) Presence and reflection in training: Support for learning to improve quality decision-making skills under time limitations. *Cyber Psychol Behav* 4(2):265–277

50. Schön DA (1984) *The reflective practitioner: How professionals think in action*. Basic Books, New York
51. Schulteis MT, Rothbaum BO (2003) Ethical issues for the use of virtual reality in the psychological sciences. *Ethical issues in clinical neuropsychology*, 243–280
52. Schwarzer R, Jerusalem M (2010) The general self-efficacy scale (GSE). *Anxiety Stress Coping* 12:329–345
53. Scoresby J, Shelton BE (2014) Reflective redo from the point of error: Implications for after action review. *Simul Gaming* 45(4–5):666–696
54. Sharot T, Delgado MR, Phelps EA (2004) How emotion enhances the feeling of remembering. *Nat Neurosci* 7(12):1376
55. Sharples S, Cobb S, Moody A et al (2008) Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29(2): 58–69. <https://doi.org/10.1016/j.displa.2007.09.005>
56. Sherman WR, Craig AB (2018) *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann, San Francisco
57. Smith S, Ericson E (2009) Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual Reality* 13(2):87–99. <https://doi.org/10.1007/s10055-009-0113-6>
58. Stajkovic AD, Luthans F (1998) Self-efficacy and work-related performance: A meta-analysis. *Psychol Bull* 124(2):240
59. Starks K (2014) Cognitive behavioral game design: a unified model for designing serious games. *Front Psychol* 5:28
60. Stauffert J, Niebling F, Latoschik ME (2020) Latency and cybersickness: Impact, causes and measures. A review. *Front Virtual Real* 1:31
61. Stein D (1998) Situated learning in adult education. In: ERIC Clearinghouse on Adult, Career, and Vocational Education, Center on Education and Training for Employment, College of Education, the Ohio State University, pp 195–202
62. United States Geological Survey (2018) Earthquake facts. Available via <https://earthquake.usgs.gov/learn/facts.php>. Accessed 4 Oct 2018

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.