Safety Science 171 (2024) 106372

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

# Virtual reality for safety training: A systematic literature review and meta-analysis

D. Scorgie [a], Z. Feng [a], D. Paes [a], F. Parisi [b], T.W. Yiu [c], R. Lovreglio [a]

[a] *School of Built Environment, Massey University, New Zealand*
[b] *Department of Electrical and Information Engineering, Politecnico di Bari, Italy*
[c] *School of Built Environment, The University of New South Wales, Sydney, Australia*

ARTICLE INFO

*Keywords:*
Virtual reality
Safety training
Meta analysis

ABSTRACT

Unsafe behaviour in the workplace and disaster events can lead to serious harm and damage. Safety training has been a widely studied topic over the past two decades. Its primary aim is to save lives and minimise damage but requires regular refreshers. New digital technologies are helping in the process of enhancing safety training for better knowledge acquisition and retention. Among them, Virtual Reality (VR) can provide an engaging and exciting training experience, and there is a need to evaluate its application and effectiveness in safety training.

This study aims to investigate VR safety training solutions applied to various industries (excluding medical and military applications), such as construction, fire, aviation, and mining. This was achieved by systematically reviewing 52 articles published between 2013 and 2021 to answer nine research questions. Fourteen domains were examined, with construction and fire safety training being the most prevalent since 2018. Findings reveal that only a small percentage (9.6 %) of the studies explicitly adopted theories while developing and testing VR applications. Additionally, this review highlights a critical need for long-term retention measurements, as only 36 % of studies provided such data. Finally, the two meta-analyses proposed in this work demonstrate that VR safety training outperforms traditional training in terms of knowledge acquisition and retention.

## 1. Introduction

Safety training is a key activity worldwide for accident prevention strategies that aim to reduce the impact of accidents and disasters on humans and property (Hale, 1984). It represents one of the keystones of many occupational health and safety programs (Robson et al., 2012). Safety training plays a key role in enhancing humans' abilities and skills to identify risks and analyse the magnitude of these risks. These abilities and skills are among the key factors determining humans' behaviour and safety while performing dangerous tasks (Sacks et al., 2013). Previous studies have identified strong evidence of the effectiveness of safety training on workers' behaviours (Robson et al., 2012). Several of them have demonstrated the significant impact of safety training on humans' disaster preparedness (Hsu et al., 2013; Nazli et al., 2014).

Traditionally, safety training has been delivered using different methods. These include safety manuals, videos, in-person or online lectures, and drills (Feng et al., 2018). However, most of these methods have critical pedagogical limitations. For example, drills often do not provide feedback to participants, which could help them assess their choices retrospectively (Gwynne et al., 2019). Moreover, traditional methods face numerous challenges in effectively illustrating what a hazardous event looks like for trainees. In most cases, replicating extreme scenarios in the physical world would be impossible, given time, cost, and safety constraints (Pedram et al., 2017).

New digital technologies are making the process of safety training more effective. Virtual Reality (VR), in particular, is one of the most promising technologies that can enhance the effectiveness of safety training. Previous studies have shown that VR is a suitable tool for designing safety training systems and investigating human behaviour in emergencies (Lovreglio, 2020). VR has been adopted in its immersive and non-immersive versions[1] to deliver safety training in many industries and for different types of disasters (Kanade and Duffy, 2022; Lovreglio, 2020). Combined with gamification, it can make the content of safety training programs more engaging and promote learning (Feng et al., 2018). To date, several works have demonstrated the benefits and limitations of VR training in different domains, such as evacuations in different disasters (Feng et al., 2018; Gagliardi et al., 2023), construction (Menzemer et al., 2023; Zhao and Lucas, 2015; Li et al., 2018; Bhoir and Esmaeili, 2015; Gao et al., 2019), and mining (Tichon and Burgess-Limerick, 2011). Broader reviews are provided by Kanade and Duffy

---

[1] Readers can refer to Section 2 for definitions.

https://doi.org/10.1016/j.ssci.2023.106372
Received 12 September 2023; Received in revised form 28 October 2023; Accepted 31 October 2023

(2022) using bibliometric analysis and Stefan et al. (2023) focusing on domains and evaluation methods. Nonetheless, there is currently a gap in the literature when it comes to a thorough review of VR applications in safety training. The need for a systematic and comprehensive understanding of how VR safety training has been implemented and tested across various domains remains evident. Combining insights from different domains can provide reliable answers on whether and the extent to which VR is a valuable safety training solution. Further, a meta-analysis comparing the effectiveness of VR safety training against traditional safety training is another research gap in the literature.

This study aims to evaluate how VR has been applied to safety training for various sectors, such as construction, fire, aviation, and mining, as well as to assess its efficacy when compared with traditional tools. In this study, the safety training under investigation is mainly about training provided to individuals to deal with a scenario or process; thus, large-scale safety programs such as civil protection and crisis management were not included. To achieve this aim, we conducted a systematic literature review assessing 52 papers published between 2013 and 2021. The knowledge gained from this review is further advanced by two meta-analyses comparing the effectiveness of VR and traditional safety training in terms of knowledge acquisition and retention. This work represents a new milestone in the VR safety training research, bringing new insights into how VR has been adopted for safety training and whether and the extent to which it is more effective than traditional training methods.

## 2. Background

### 2.1. Virtual reality

The modern-day concept of VR originated from the development of the stereoscope in the 19th century. Modern VR technology had its initial introduction with the development of the Sensorama device in the 1960s, which was designed to incorporate five short films, engaging multiple senses, including vision, hearing, olfaction, and haptics, creating an immersive experience for the users (Rendevski et al., 2022). In 1994, the "3i" features of VR were proposed, namely, Immersion, Interaction, and Imagination (Burdea and Coiffet, 2003). They allow users to experience the virtual world beyond reality (Berg and Vance, 2017). In the following decades, VR has been demonstrated as a powerful tool and applied in various industries for different purposes (LaValle, 2023).

Some VR setups deliver immersive experiences, while others deliver non-immersive ones (Buttussi and Chittaro, 2021; Paes et al., 2021). Both types of VR are computer-generated simulations. While immersive VR involves the use of visual display devices delivering a high level of immersion and visual realism, closely resembling real-life experiences, non-immersive VR is often based on standard computers (Jennett et al., 2008; Lee and Lee, 2004; Robertson et al., 1993).

The most well-known immersive VR devices are VR headsets (head-mounted displays, HMDs) and projection-based displays. VR headsets are the most common visual display devices for VR that can deliver the 3i features, and their market size is growing rapidly (Chen et al., 2021). These headsets provide users with a stereoscopic 3D visual experience and often incorporate head-tracking technology to allow users to look around and interact with the virtual world. HMDs are typically lightweight and designed for comfortable wear, making them ideal for extended VR experiences (Ito et al., 2021).

There are several types of head-mounted displays for VR, each offering different features and capabilities (Angelov et al., 2020). Tethered HMDs are physically connected (usually wired) to a computer and provide high-quality graphics and performance. They typically offer an immersive experience, but their physical connections limit users' freedom of movement. In some cases, these limitations can be overcome using wireless communication with a computer, which could often be unstable. In turn, standalone HMDs feature built-in hardware required to fully reproduce a computer-generated virtual environment, such as graphics processing units, display for image visualisation, measurement sensors, headphones, storage, and batteries. Despite not needing an external computer, they trade off freedom of movement with performance related to hardware specifications. The smartphone-based VR headsets, named mobile HMDs, produce the VR content relying on the user's smartphone. These headsets feature a slot to place the smartphone that acts as the processing unit and visual display for the VR experience. Despite being affordable and widely accessible, they perform poorly than standalone or tethered HMDs. Among the most common models are the Sony PlayStation, Oculus Rift, HTC Vive Pro, HTC Vive Cosmos, Valve Index and Samsung HMD Odyssey+ (Angelov et al., 2020).

Projection-based VR provides users with an interactive and immersive experience by projecting virtual content onto screens or surfaces. Among the most widespread setups is the Cave Automatic Virtual Environment (CAVE) (Creagh, 2003; Sandin et al., 1993), which displays 3D images on walls, floors, and ceilings. Users wear special 3D glasses to perceive the 3D images and interact with the virtual environment using handheld controllers. CAVEs can deliver high-level visuals and immersion but come with limitations, such as the cost and complexity of the setup and limited freedom of movement (Havig et al., 2011). Powerwalls, in turn, are basically large-scale, high-resolution displays also used to create interactive virtual environments on flat or curved screens. It shares many features with the CAVE system, such as the use of 3D glasses and devices to interact with the virtual environment while trading off immersion for cost and mobility (de Vasconcelos et al., 2019).

In non-immersive VR, users do not undergo a fully multisensorial simulation, but their perception of the physical world is not blocked so that they maintain a certain degree of awareness of the real environment. These technologies are desktop-based: the virtual experience content is accessed on different kinds of screens, while the interaction is performed using a mouse, keyboard, or other input devices. These technologies are typically more accessible and do not require specialised hardware like HMDs.

### 2.2. Safety training

Unsafe behaviours can potentially lead to fatalities, injuries, or economic repercussions in the context of workplace accidents or disaster events. Safety training stands as a primary mechanism for enhancing human behaviour while engaging in dangerous tasks (Sacks et al., 2013). It has become a globally vital factor of accident prevention strategies, aiming to minimise the impact of accidents and disasters on humans and property (Hale, 1984). It is also a legal requirement in numerous countries across various industries, underscoring its status as a regulatory imperative (Loosemore and Malouf, 2019). For instance, in New Zealand, construction workers are required to attend a Health and Safety course every two years (Site Safe Passport Courses - Third Edition, 2019), and in the US, employers take significant responsibilities regarding worker safety training requisites (Weiss, 2022). Also, safety training related to disasters has become a legal requirement worldwide, such as Australia's annual evacuation exercise for emergencies in facilities and the UK annual fire drill for workplaces, including schools (Fire safety in the workplace, Model Work Health and Safety Regulations, 2023).

Traditionally, safety training methods consist of videos, lecture slides and safety manuals, which tend to create a passive learning environment for trainees (Buttussi and Chittaro, 2021; Avveduto et al., 2017). They have been criticised for their potential shortcomings in facilitating knowledge acquisition, which could be attributed to the limited sense of presence and engagement they offer. Consequently, these methods are argued to be insufficiently motivating, leading to reduced trainee attentiveness and concentration over short durations (Cherrett et al., 2009; Harfield et al., 2007). Additionally, they have been associated with high costs and may not represent the most optimal solution for effective safety training (Forst et al., 2013).

Within the transition to industrial 4.0, the utilisation of digital technology in safety training is becoming more and more widespread and applied across various industries (Liang et al., 2020; Validakis, 2015). For instance, U.S. military and intelligence community have used VR for many safety training projects, such as nuclear attacks by terrorists and urban warfighting operations (Wilson, 2008), and the British Safety Council used Augmented Reality (AR) to deliver occupational health and safety knowledge (Virtual and augmented reality training). Previous studies have also proved that innovative safety training methods are overall better than traditional methods (Feng et al., 2021; Lovreglio et al., 2021).

## 3. Methodology

This study incorporates a systematic literature review and two meta-analyses. The systematic literature review employed the “five steps for conducting a systematic review” by Khan et al. (2003): defining research questions, identifying the relevant work, assessing the quality of the studies, summarising the evidence, and interpreting the findings. Two separate meta-analyses have been conducted to assess and compare knowledge acquisition in the context of VR and traditional training, as well as to compare knowledge retention between the two training solutions. Hedge’s g is employed to assess the standard mean differences in the two meta-analyses, as it is the best solution when the sample sizes are small (<20) (Shadish et al., 2014; Takeshima et al., 2014).

### 3.1. Framing the research questions

To comprehensively understand how VR is used for safety training, we expand upon the examination of the state of research by considering the types of technique, the topics taught, the methods to evaluate learning outcomes, and the delivered learning outcomes (Benitti, 2012). To address these aspects, nine research questions were formulated as follows.

Question 1: What domains and topics have been taught in VR safety training?
Question 2: Which theories were used to develop and test the VR safety training?
Question 3: Which types of VR hardware have been used?
Question 4: Which measurements were used to assess the training effectiveness?
Question 5: What are the learning outcomes?
Question 6: What are the instructional methods?
Question 7: What are the experimental results?
Question 8: What are the outcomes of previous studies comparing VR training and traditional training?
Question 9: What are the outcomes of previous studies comparing different VR solutions?

### 3.2. Identify the relevant works

The publications included in the systematic literature review were collected from Scopus and Web of Science, two world-leading scholarly databases. Scopus is the largest scientific literature database (Bar-Ilan, 2008), and Web of Science is another large database of peer-reviewed literature in science and engineering (Wuchty et al., 2007).

The literature employs inconsistent terminology. For instance, different papers refer to virtual reality as VR, virtual environment, or virtual simulation. To achieve the broadest possible range of publication coverage, the search string used is as follows: “virtual reality” OR “virtual environment” OR “virtual simulation” OR “VR” AND training OR education AND safety. The searches yielded 1688 results, 867 from Scopus and 821 from Web of Science. Upon organising the identified results by title, 551 duplicates within the set were found. After removing the duplicates, the pool of eligible papers was constrained to those composed in English and published from 2012 to 2021. The items were limited to non-military and non-medical-related topics. The reason is that most VR medical studies mainly focus on the safety implications of administering medication or treatments to patients instead of providing safety training to individuals to deal with a scenario or process. Therefore, those medical studies fall beyond this research’s scope. Similarly, military studies were excluded as they focus on non-civilian applications, which are also out of the scope of this research.

The titles and abstracts of each paper were examined; only the ones containing VR-based safety training and related terms were regarded as eligible. Following the screening, 302 papers remained. 193 of these were excluded via quickly reading the full text according to the following criteria: (i) there was no VR training prototype proposed; (ii) there were no following no experiments or case studies; (iii) there was no data analysis of the outcome to evaluate and validate the prototype. The remaining 109 papers were selected. Fig. 1 shows the selection process.

### 3.3. Assessing the quality of studies

To evaluate the relevance of the 109 selected articles to answer the nine research questions, a 3-point Likert scale (3 – high, 2- medium and 1 – low) was adopted to assess three dimensions of each paper: (i) how relevant is the topic or focus of the paper to “VR safety training”; (ii) the extent to which the prototype and design align with the questions of this review; and (iii) how appropriate and relevant the research results are to the questions of this review. The scoring criteria were developed based on Connolly’s quality assessment approach (Connolly et al., 2012). Each paper could get the highest point of 9 and the lowest point of 3 in all three dimensions. As a result, 17 papers were marked as 8-point, 21 papers were marked as 7-point, 14 papers were marked as 6-point, 33 papers were marked as 5-point, 19 papers were marked as 4-point, and 5 papers were marked as 3-point. Ultimately, 52 papers which were scored 6 points or above were shortlisted and adopted in this review. These papers (called hereafter the *eligible papers*) were summarised in Table 1.

### 3.4. Meta analysis tools

One of the sub-goals of this work is to run a comparison between the efficacy of VR training and traditional training using two meta-analyses. Meta-analyses are tools allowing researchers to formally and systematically gather data available in specific research domains and draw conclusions based on all currently published information (Rosenthal and Schisterman, 2010). In this study, quantitative comparison was done by focusing on knowledge acquisition and retention. To date, different solutions have been adopted to assess the effect size for two independent groups (Becker, 2000). As shown in Section 4.8, different scales have been used to quantify safety knowledge. A meta-analysis was carried out only using standardised mean differences (i.e., the mean difference expressed in SD units) instead of simple mean differences. Cohen’s d and Hedge’s g are among the most used matrixes in the literature. In this study, Hedge’s g was used to assess the effect size as it represents the best solution when the sample sizes are small (<20) (Andrade, 2020; Takeshima et al., 2014; Shadish et al., 2014). Its formula is shown in Equation (1).

$$g = \frac{M_1 - M_2}{SD^*_{pooled}} \tag{1}$$

where$M_1$ and $M_2$ are the means for groups 1and 2 and $SD^*_{pooled}$ is defined as in Equation (2).

$$SD^*_{pooled} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}} \tag{2}$$

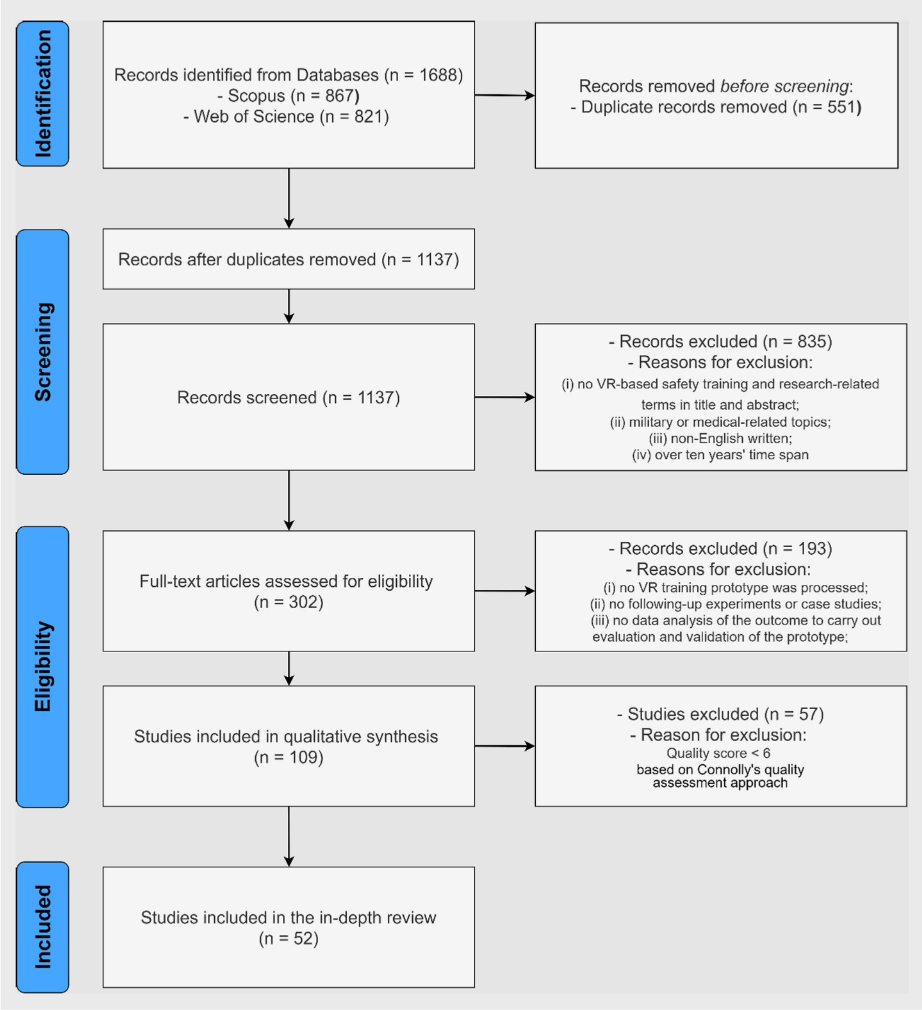


**Fig. 1.** Literature selection process.

**Table 1**
Papers quality scores according to Connolly's quality assessment approach.

| 6-point score papers | 7-point score papers | 8-point score papers |
|---|---|---|
| (Chittaro et al., 2014) (Dzeng et al., 2015) (Grabowski and Jankowski, 2015) (Jeon, 2016) (Schwebel et al., 2016) (Pham et al., 2018a) (Asghar et al., 2019) (Wang and Yang, 2019) (Kim et al., 2020) (Lacko, 2020) (Rahmalan et al., 2020) (Afzal and Shafiq, 2021) (Bhagwat et al., 2021) (Satapanasatien et al., 2021) | (Kinateder et al., 2013) (Sacks et al., 2013) (Burigat and Chittaro, 2016) (Prendinger et al., 2016) (Maillot et al., 2017) (Jung and Ahn, 2018) (Lu and Davis, 2018) (Pham et al., 2018b) (Liang et al., 2019) (Oliva et al., 2019) (Saeidi et al., 2019) (Vahdatikhaki et al., 2019) (Eiris et al., 2020) (Saghafian et al., 2020) (Dhalmahapatra et al., 2021) (Khan et al., 2021) (Paszkiewicz et al., 2021) (Poyade et al., 2021) (Rahouti et al., 2021) (Valentine et al., 2021) | (Chittaro and Buttussi, 2015) (Avveduto et al., 2017) (Li et al., 2017) (Buttussi and Chittaro, 2018) (Schwebel et al., 2018) (Çakiroğlu and Gökoğlu, 2019) (Makransky et al., 2019) (Feng et al., 2020) (Jeelani et al., 2020) (Nykänen et al., 2020) (Pedram et al., 2020) (Buttussi and Chittaro, 2021) (Jacobsen et al., 2021) (Lovreglio et al., 2021) (Morélot et al., 2021) (Shiradkar et al., 2021) |

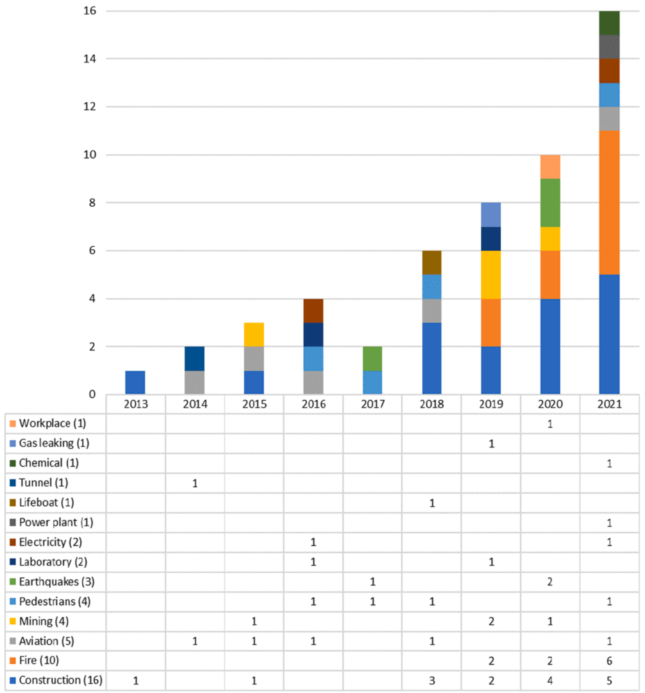
where $n_1$ and $n_2$ are the sample sizes for group (1) and (2) while $SD_1$ and $SD_2$ are the standard deviations for groups 1and 2.

## 4. Results

The eligible papers were submitted for coding and analysis using a data extraction spreadsheet. A summary of all the papers reviewed in this work is provided as supplementary material.

### *4.1. Domains and topics*

Since 2013, an increasing trend can be observed in terms of implementing VR for safety training in various domains (Fig. 2). Fourteen domains were covered by VR safety training, with the majority being construction safety training (16 papers) and fire safety training (10 papers). The construction and fire domains began to gain more attention in 2018. A few studies have focused on the aviation, mining, pedestrians, and earthquake domains, while some studies developed and tested VR for laboratory, electricity, lifeboat, tunnel, chemical, gas-leaking, and workplace safety training. Also, the evolution of VR safety training domains over time has exhibited both expansion and diversification. This evolution is exemplified by the transition from a single domain in 2013 to the inclusion of six domains by 2021.



| | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|---|---|---|---|---|---|---|---|---|---|
| ■ Workplace (1) | | | | | | | | 1 | |
| ■ Gas leaking (1) | | | | | | | 1 | | |
| ■ Chemical (1) | | | | | | | | | 1 |
| ■ Tunnel (1) | | 1 | | | | | | | |
| ■ Lifeboat (1) | | | | | | 1 | | | |
| ■ Power plant (1) | | | | | | | | | 1 |
| ■ Electricity (2) | | | | 1 | | | | | 1 |
| ■ Laboratory (2) | | | | 1 | | | 1 | | |
| ■ Earthquakes (3) | | | | | 1 | | | 2 | |
| ■ Pedestrians (4) | | | | 1 | 1 | 1 | | | 1 |
| ■ Mining (4) | | | 1 | | | | 2 | 1 | |
| ■ Aviation (5) | | 1 | 1 | 1 | | 1 | | | 1 |
| ■ Fire (10) | | | | | | | 2 | 2 | 6 |
| ■ Construction (16) | 1 | | 1 | | | 3 | 2 | 4 | 5 |

**Fig. 2.** VR safety training domains and publication years.

### 4.2. *Adopted theories*

A few studies explicitly demonstrated the theories they used to develop and test their VR training systems. (Avveduto et al., 2017) applied the Kirkpatrick model (i.e., levels of training evaluation) to evaluate their VR safety training application. The Kirkpatrick model measures a training program from four levels: reaction, learning, behaviour, and results. (Çakiroğlu and Gökoğlu, 2019; Kirkpatrick, 2015) followed the Behavioural Skills Training (BST) model to provide safety training. The BST model suggests an active learning environment consisting of instruction, modelling, rehearsal, and feedback (Himle and Miltenberger, 2004; Stewart et al., 2007). Makransky et al. (2019) investigated the motivational and cognitive benefits of training in VR. They proposed hypotheses and developed an assessment framework based on the interest theory (Renninger and Hidi, 2015), self-determination theory (Ryan and Deci, 2016), cognitive theory of multimedia learning (Mayer, 2005), cognitive load theory (Sweller, 2011), and the embodied cognitive theory (Wilson, 2002). These theories lay the foundation for assessing and understanding motivation and cognition in VR safety training (Nykänen et al., 2020). developed an evaluation framework for VR safety training utilising the social cognitive theory (Bandura et al., 1999) and the theory of planned behaviour (Ajzen, 1991). Based on these theories, (Nykänen et al., 2020) identified factors that are crucial in safety performance and safety outcomes, such as self-efficacy, locus of control, outcome expectancies, safety knowledge, and safety motivation. (Satapanasatien et al., 2021) adopted the game-based learning (GBL) theory (Plass et al., 2015; Granic et al., 2014) to guide their development of a VR fire safety training system. The GBL theory suggests the following contributing factors to an effective learning environment: motivation, player engagement, adaptation, and graceful failure.

### 4.3. *Types of VR hardware*

The type of VR hardware was presented into three distinct categories: headset, screen-based, and projection-based. Thirty-two studies employed VR headsets. Twenty studies applied various screen-based VR technologies, such as multi-screen, curved-screen, and customised screens (Jeon, 2016; Khan et al., 2021; Maillot et al., 2017). Projection-based VR is classified into two types: stereo and non-stereo. Stereo-projection-based VR was utilised in two studies (Pedram et al., 2020; Sacks et al., 2013). Pedram et al., (2020) employed a 360 VR, which was a 10 m diameter, 4 m high cylindrical screen that displayed a stereo 3D 360-degree virtual environment, while Sacks et al. (2013) used a power wall, which comprised three rear-projection screens, each measuring 2.4 m in width and 1.8 m in height, placed in a theatrical arrangement with a 150-degree separation between adjacent screens. Three studies employed non-stereo projection-based VR: two pedestrian safety experiments and one fire extinguisher operation experiment (Khan et al., 2021; Maillot et al., 2017; Morélot et al., 2021). Fig. 3 illustrates the growth trend in adopting different types of hardware, with a particular emphasis on VR headset usage, which experienced a substantial increase between 2019 and 2021 compared to the preceding years.

### 4.4. *Effectiveness measurements*

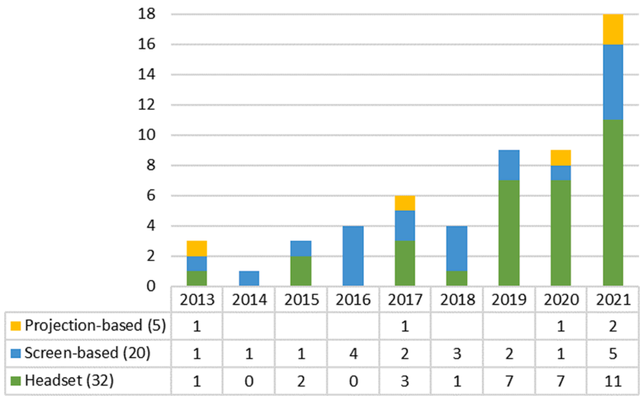
#### 4.4.1. *Types of measurements*

General measurements consisted of knowledge acquisition, user experience and retention. Most studies adopted knowledge acquisition as a measurement to assess training effectiveness. Questionnaire surveys, verbal descriptions, in-situ performance and observation by the trainer were the common data collection tools. User experience consists of various factors depending on the aim and objectives of the individual study, such as engagement, presence, enjoyment, ease of use, navigation, and sense of control. Most of the user experiences were collected by a Likert Scale survey. Some studies employed the existing surveys for the specific experience. For instance, Morélot et al. (2021) and Kim et al. (2020) used Slater-Usoh and Steed to measure the sense of presence; Buttussi and Chittaro (2021) and Poyade et al. (2021) adopted the System Usability Scale to measure the usability of the training prototypes. Retention is a relevant new measurement focusing on knowledge in most studies. Eighteen studies involved knowledge retention, and thirteen of them were published between 2017 and 2021, and five were published between 2011 and 2016. Also, all of them measured short-term knowledge retention. In contrast, one study (Grabowski and Jankowski, 2015) collected three months of knowledge retention data, and the rest of the retention tests were conducted within four weeks.

#### 4.4.2. *Measurement time*

Most studies examining knowledge acquisition typically employed a pre-test and an immediate post-test to assess the difference in knowledge between these two assessments (Feng et al., 2021; Jung and Ahn, 2018; Rahouti et al., 2021). Conversely, studies focused on user experiences mainly adopted a sole post-test evaluation (Asghar et al., 2019; Wang and Yang, 2019).

Nineteen papers were subject to retention tests, and as depicted in Fig. 4, nearly half of the papers exhibited a retention duration ranging



| | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|---|---|---|---|---|---|---|---|---|---|
| ■ Projection-based (5) | 1 | | | | 1 | | | 1 | 2 |
| ■ Screen-based (20) | 1 | 1 | 1 | 4 | 2 | 3 | 2 | 1 | 5 |
| ■ Headset (32) | 1 | 0 | 2 | 0 | 3 | 1 | 7 | 7 | 11 |

**Fig. 3.** Hardware type by year.

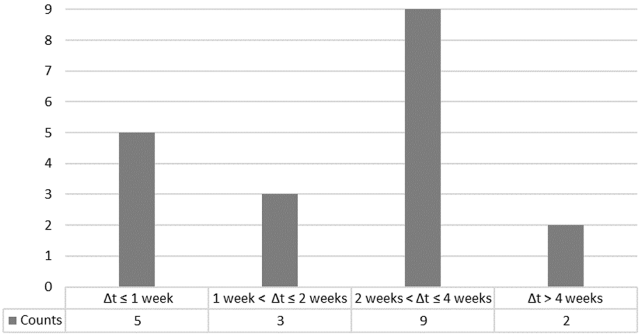


**Fig. 4.** Retention measurement timing (ΔT).

from two to four weeks. It was found that only two papers published in earlier years extended their retention measurements beyond four weeks, with one study reporting a retention interval of three months and another measuring retention over one year (Grabowski and Jankowski, 2015; Kinateder et al., 2013). Additionally, only one study incorporated two retention assessments, with the first assessment conducted after one week and the second after one year (Kinateder et al., 2013).

### 4.5. *Learning outcomes*

Hazard recognition and management, safety actions, and proper utilization of equipment and tools are identified as the main training outcomes, each including various specific tasks in individual papers. More than 50 % of the papers related to construction (n = 16) chose hazard identification as the primary training outcome, which is often associated with the prevention of hazards, while certain studies solely concentrate on the nature of hazards, including their causes, types, and identification (Pham et al., 2018b). Some studies focused on the procedures of safety training, which aimed to inculcate appropriate techniques for movement, operation, and installation (Kim et al., 2020; Valentine et al., 2021). Nonetheless, only a few studies encompassed multifaceted training outcomes, combining hazard identification and procedural training. In the case of fire-related papers (n = 20), 50 % of the studies used fire extinguishers as a case study for training purposes. These studies focused on training individuals on techniques such as PASS, the classification of fires, the types of extinguishers, and the distance from the fire when operating the extinguisher. Additionally, seven papers were dedicated to training individuals in safety behaviour skills, such as basic fire reaction knowledge and identifying the nearest exit. All papers related to aviation implemented safety training procedures that focused on best practices for escaping the cabin after an aircraft incident, such as opening the overwing exits. All experiments related to mining had a primary focus on safety behaviour, yet each experiment emphasised specific trained behaviours, such as the appropriate actions required for blasting operations (Grabowski and Jankowski, 2015). The pedestrian experiments were consistently trained in safe road-crossing practices. The remaining papers consistently centered on the three predetermined outcomes and incorporated specific tasks customised to suit the pertinent domain and topic.

### 4.6. *Instructional method*

Approximately half of the studies involved instruction methods. According to the feedback time, the following three instructional methods were summarised as the most used:

(i) Warning when the user tends to make the incorrect move or not make the movement for a specific time;
(ii) Immediate feedback with texture recommendation and/or verbal description straight after the users make a mistake;
(iii) Delayed feedback with text or oral explanatory (or both) after the training.

Several studies implemented a verbal or textual warning mechanism before a mistake is made, particularly when the user tends to execute an incorrect action or fails to perform an action within a designated timeframe (Lu and Davis, 2018; Schwebel et al., 2018).

Immediately feedback for action, regardless of a consequence, was a standard instructional method, particularly in response to erroneous action, and has been widely employed. The most prevalent mode of feedback provision is via verbal description accompanied by a highlighted test message that indicates the correct response (Çakiroğlu and Gökoğlu, 2019). Specific experiments have exclusively employed textual messages to correct irreversible errors with a few seconds pause (Buttussi and Chittaro, 2018; Chittaro and Buttussi, 2015). Few other customised feedback modes were used, such as verbal warning or text warning after the erroneous action without providing the correct answer; red and green flashlights present the incorrect and correct movement, respectively (Feng et al., 2020; Pham et al., 2018b). Additionally, some studies have employed feedback delivery after training, involving verbal and textual descriptions of the erroneous and correct actions (Paszkiewicz et al., 2021; Schwebel et al., 2016).

### 4.7. *Experiment results*

Fifteen papers only focused on the assessment of VR training prototypes. Most of the studies reported positive results of VR training, although a few studies argued that VR training is average (Schwebel et al., 2018). Some studies investigated how different factors affect specific training outcomes or user experiences. For instance, (Xueqing and Steven, 2018) indicated that sound and priming factors could affect cognition, with priming factors having a more significant effect on decision-making than sound. (Chittaro et al., 2014) found that fear and time spent could help to improve knowledge retention, while (Saghafian et al., 2020) suggested that emotional and bodily experiences are good supplements for training.

Twenty-four papers investigated the differences between VR and traditional training methods, fourteen papers compared different types of VR training, and four did both. Over half of the papers showed that the level of immersion affects knowledge acquisition and user experiences. Higher levels of immersion were associated with better knowledge acquisition and user experiences (Buttussi and Chittaro, 2021; Makransky et al., 2019; Morélot et al., 2021). However, the level of immersion does not always play a positive role in knowledge acquisition or user experiences. (Jung and Ahn, 2018) argued that desktop VR was better for procedural knowledge than HMD, Poyade et al. (2021) and Satapanasatien et al. (2021) provided evidence to show that traditional training methods have equal effectiveness in both knowledge acquisition and user experience.

### 4.8. *Comparison of VR training and traditional training*

#### 4.8.1. *Qualitative comparison*

Half of the studies examined the distinction between virtual reality (VR) and traditional training methods, which mainly encompassed video, paper-based material, and lecture-based material, while two studies adopted on-site training and observation as the traditional training method. Findings indicate that VR-based training was generally superior to traditional methods regarding knowledge acquisition, user experience, and knowledge retention. In accordance with various measures, VR-based training has been observed to surpass traditional methods 47 times, whereas an equivalence in effectiveness between VR-based and traditional methods was established in 13 instances. Only one study indicates the superiority of traditional methods over VR-based training. Table 2 focuses on the knowledge acquisition and retention of VR and three main traditional methods.

**Table 2**
List of studies comparing VR and different traditional training methods (+ : VR performs better than the traditional; - : VR performs worse than the traditional; = : no difference).

| Article | VR and traditional training solutions | Knowledge acquisition | Knowledge retention |
|---|---|---|---|
| (Avveduto et al., 2017) | VR vs. video | +/=* | + |
| (Li et al., 2017) | VR vs. video | + | + |
| (Lovreglio et al., 2021) | VR vs. video | + | + |
| (Liang et al., 2019) | VR vs. video | + | + |
| (Poyade et al., 2021) | VR vs. video | = | N/A |
| (Jeon, 2016) | VR vs. video | + | N/A |
| (Lacko, 2020) | VR vs. video | + | + |
| (Buttussi and Chittaro, 2021) | VR vs. paper-based | + | + |
| (Chittaro and Buttussi, 2015) | VR vs. paper-based | = | + |
| (Li et al., 2017) | VR vs. paper-based | + | + |
| (Makransky et al., 2019) | VR vs. paper-based | = | = |
| (Burigat and Chittaro, 2016) | VR vs. paper-based | +/=* | N/A |
| (Kinateder et al., 2013) | VR vs. paper-based | + | = |
| (Satapanasatien et al., 2021) | VR vs. paper-based | + | N/A |
| (Morélot et al., 2021) | VR vs. lecture-based | +/=* | N/A |
| (Nykänen et al., 2020) | VR vs. lecture-based | + | + |
| (Sayli et al., 2021) | VR vs. lecture-based | = | + |
| (Jung and Ahn, 2018) | VR vs. lecture-based | + | N/A |
| (Rahouti et al., 2021) | VR vs. lecture-based | + | + |
| (Sacks et al., 2013) | VR vs. lecture-based | + | + |
| (Bhagwat et al., 2021) | VR vs. lecture-based | N/A | N/A |
| (Dzeng et al., 2015) | VR vs. lecture-based | = | N/A |

*Different types of VR showed different results when compared to traditional methods.

For what concerns knowledge acquisition and retention, the findings indicate that, for the most part, VR outperforms traditional methods in both domains. Specifically, most of the studies reviewed in the analysis report that VR is superior to traditional methods in knowledge acquisition. However, two studies demonstrate equivalent outcomes between VR and each traditional method, respectively. Only two studies suggest comparable knowledge retention outcomes between VR and paper-based approaches. In contrast, video and lecture-based training exhibit less efficacy than VR. Further, some studies included in Table 2 provide evidence on whether VR performs better or worse than traditional methods in knowledge acquisition and retention. In these studies, the comparison was carried out based on other aspects. For instance, two studies only compared how participants performed safety tasks after the training instead of measuring participant knowledge (Bhagwat et al., 2021). Another study compared the different training only in terms of user experience (Bhagwat et al., 2021).

To provide a quantitative comparison of the effectiveness of VR training in terms of knowledge acquisition and knowledge retention, two meta-analysis are carried out in the following sections using 11 studies listed in Table 2. Some studies were excluded as they did not provide all the necessary data required for the two meta-analyses (i.e. means, standard deviations, and sample size of the groups).

Twelve papers were reviewed that compared different VR training devices, which we have classified into three types: headset-based, projection-based, and screen-based. In terms of knowledge acquisition, the studies indicated that headset-based VR was equal to or better than screen-based VR, while projection-based VR was shown to be better than screen-based VR. However, there were no studies that directly compared headset-based and projection-based VR. Regarding user experience, the studies showed that headset-based VR was equal to or better than screen-based VR, while projection-based VR was equal to screen-based VR. Again, no studies compared the user experience of headset-based and projection-based VR.

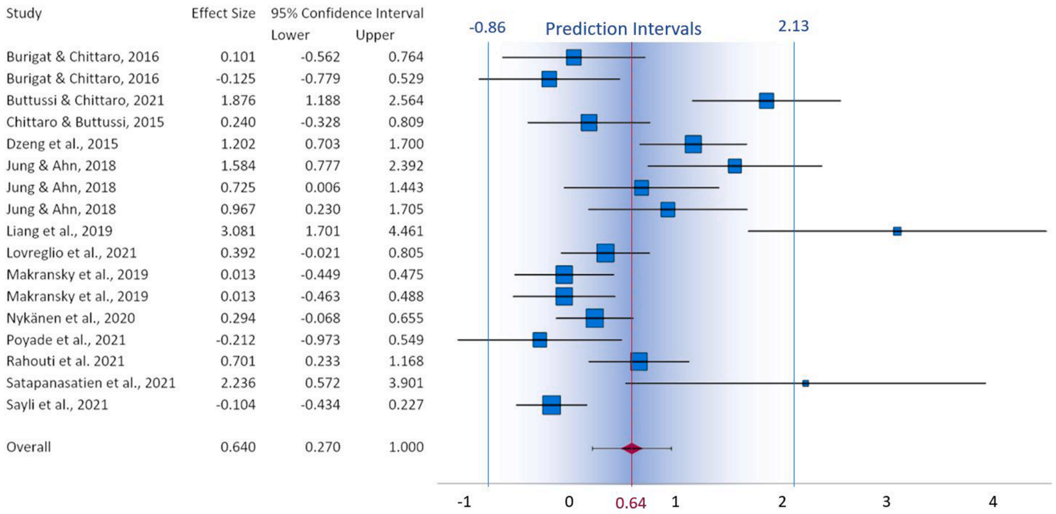
### 4.8.2. Meta-analysis

Two meta-analyses are proposed in this paper. The first one aims to compare the knowledge acquisition between VR and traditional training solutions, while the second analysis compares both solutions regarding knowledge retention. Knowledge acquisition and retention are measured using different scales in the studies used for this analysis. This lack of consistency is due to the fact that these studies focused on different safety domains and learning objectives, as explained in Section 4.1. Further, given the limited sample size for both analyses, we used Hedge's g for the analysis. The proposed analyses used mixed effect models in SPSS v.29.

The Forest Plot of the meta-analysis on knowledge acquisition is reported in Fig. 5. This analysis includes 13 studies. Since some of these studies compared different VR setups with a single traditional setup, they provide multiple values for the analysis. Finally, some of the original publications did not provide all the data required for the analysis, and the authors of these works provided supplemental material under requests sent by email. These studies are marked in Fig. 5 with the * symbol. The Forest Plot in Fig. 5 shows the overall effect size of the mixed-effect model. The effect size measured using the Hedge's g is equal to 0.640 and is statistically different from zero (p-value $< 0.001$). According to the literature, this value indicates a medium-large effect (Durlak, 2009). This result indicates that VR training performs better than traditional training in terms of knowledge acquisition. In other words, participants were more likely to learn more when using VR training. Given an effect size of 0.95, we estimate the prediction intervals assessing the true effect size are $-0.86$ and 2.13. These intervals are calculated using the equations in (Borenstein et al., 2017). This result indicates that there might still be a 19 % probability in which participants might acquire more knowledge using traditional training.

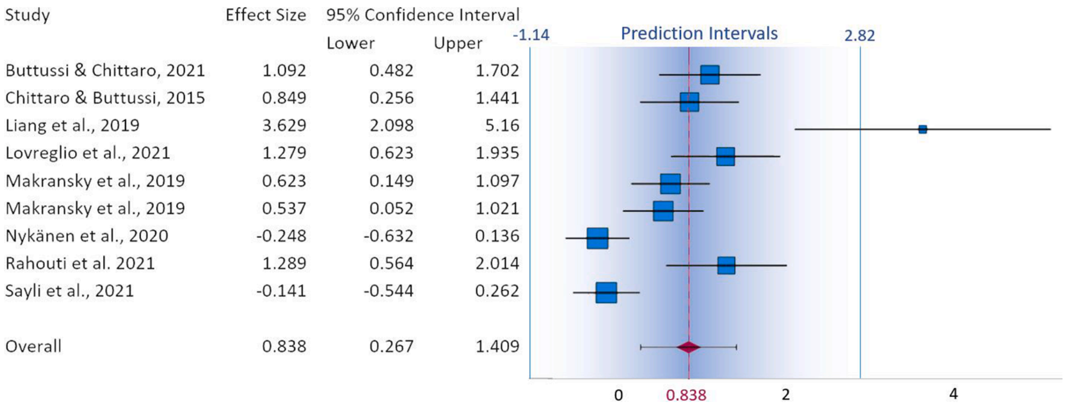
The homogeneity of the individual effect sizes reported in Fig. 5 is also calculated in this analysis. The results ($Q(df = 16) = 80.2$, p-value $< 0.001$, $I^2 = 85.6$ %) show a high heterogeneity (Higgins et al., 2003). These results indicate that other factors might explain the different effect sizes observed in the individual studies used for this analysis. As such, future analysis is required to assess the impact of different moderator effects. This was not possible in this study, given the limited number of studies available for the analysis. Finally, Egger's Regression Test was used to assess the publication bias of this analysis. The results of this test (p-value $= 0.110$) indicate that the data does not show evidence of publication bias.

The Forest Plot of the meta-analysis on knowledge retention is reported in Fig. 6. This analysis only includes eight studies, as fewer studies investigated the long-term impact of VR training on participants' knowledge. One study compared two different VR setups with a single traditional setup (Makransky et al., 2019); as such, this study provides two values. Finally, one of the original publications (Lovreglio et al., 2021) did not provide all the data required for the analysis, and the authors of these works provided supplemental material under a request sent by email. This study is marked in Fig. 5 with the * symbol. The Forest Plot in Fig. 6 shows the overall effect size of the estimated mixed-effect model. The effect size (i.e., Hedge's g) is equal to 0.838, and it is statistically different from zero (p-value $< 0.001$). This value indicates a significant effect (Durlak, 2009). This result indicates that VR training performs better than traditional training in terms of knowledge retention. In other words, participants were more likely to remember more safety information when using VR training. Given an effect size of 0.95, we estimate the prediction intervals assessing the true effect size are



| Study | Effect Size | 95% Confidence Interval Lower | Upper |
|---|---|---|---|
| Burigat & Chittaro, 2016 | 0.101 | -0.562 | 0.764 |
| Burigat & Chittaro, 2016 | -0.125 | -0.779 | 0.529 |
| Buttussi & Chittaro, 2021 | 1.876 | 1.188 | 2.564 |
| Chittaro & Buttussi, 2015 | 0.240 | -0.328 | 0.809 |
| Dzeng et al., 2015 | 1.202 | 0.703 | 1.700 |
| Jung & Ahn, 2018 | 1.584 | 0.777 | 2.392 |
| Jung & Ahn, 2018 | 0.725 | 0.006 | 1.443 |
| Jung & Ahn, 2018 | 0.967 | 0.230 | 1.705 |
| Liang et al., 2019 | 3.081 | 1.701 | 4.461 |
| Lovreglio et al., 2021 | 0.392 | -0.021 | 0.805 |
| Makransky et al., 2019 | 0.013 | -0.449 | 0.475 |
| Makransky et al., 2019 | 0.013 | -0.463 | 0.488 |
| Nykänen et al., 2020 | 0.294 | -0.068 | 0.655 |
| Poyade et al., 2021 | -0.212 | -0.973 | 0.549 |
| Rahouti et al. 2021 | 0.701 | 0.233 | 1.168 |
| Satapanasatien et al., 2021 | 2.236 | 0.572 | 3.901 |
| Sayli et al., 2021 | -0.104 | -0.434 | 0.227 |
| Overall | 0.640 | 0.270 | 1.000 |

**Fig. 5.** Forest Plot of the random effect model comparing the knowledge acquisition between VR training and Traditional training (note: some studies are entered multiple times in the analysis as they compare different VR setups with a single traditional setup).



| Study | Effect Size | 95% Confidence Interval Lower | Upper |
|---|---|---|---|
| Buttussi & Chittaro, 2021 | 1.092 | 0.482 | 1.702 |
| Chittaro & Buttussi, 2015 | 0.849 | 0.256 | 1.441 |
| Liang et al., 2019 | 3.629 | 2.098 | 5.16 |
| Lovreglio et al., 2021 | 1.279 | 0.623 | 1.935 |
| Makransky et al., 2019 | 0.623 | 0.149 | 1.097 |
| Makransky et al., 2019 | 0.537 | 0.052 | 1.021 |
| Nykänen et al., 2020 | -0.248 | -0.632 | 0.136 |
| Rahouti et al. 2021 | 1.289 | 0.564 | 2.014 |
| Sayli et al., 2021 | -0.141 | -0.544 | 0.262 |
| Overall | 0.838 | 0.267 | 1.409 |

**Fig. 6.** Forest Plot of the random effect model comparing the knowledge retention between VR training and Traditional training (note: some studies are entered multiple time in the analysis as they compare different VR setups with a single traditional setup).

−1.14 and 2.82 (these intervals are calculated using the equations in (Borenstein et al., 2017)). This result indicates that there might still be 18 % probability in which participants might retain more knowledge using traditional training.

The homogeneity of the individual effect sizes reported in Fig. 5 is also calculated in this second meta-analysis. The results (Q (df = 8) = 55.4, p-value < 0.001, $I^2 = 89.5$ %) show, once again, a high heterogeneity (Higgins et al., 2003). As such, the results indicate that there might be some other factors that can explain the different effect sizes observed in the individual studies used for this analysis. One of these factors could be the retention time, as this was different in the studies included in this analysis, as shown in Fig. 4. Future analysis is required to assess the impact of different moderator effects. This was not possible in this study, given the limited number of studies available for the analysis. Finally, Egger's Regression Test was used to assess the publication bias of this analysis. The results of this test (p-value = 0.001) indicate that the data show evidence of publication bias. This bias can be explained by the fact that all the selected studies come from peer-review journals, and it can be explained by the tendency of journals to reject studies that do not show statistical differences (Thornton and Lee, 2000). This bias could be solved in future studies by also including studies that were not published in peer-reviewed journals like white papers and pre-prints.

### *4.9. Comparison of different types of VR training devices*

Seventeen papers were reviewed that compared different VR training devices, which we have classified into three types: headset-based, projection-based, and screen-based. The results are reported in Table 3. In terms of knowledge acquisition, the studies indicated that headset-based VR was equal to or better than screen-based VR, while projection-based VR was shown to be better than screen-based VR. However, there were no studies that directly compared headset-based and projection-based VR. Regarding user experience, the studies showed that headset-based VR was equal to or better than screen-based VR, while projection-based VR was equal to screen-based VR. Once again, no study was found to compare the user experience of headset-based and projection-based VR. In this instance, it was not possible to run a quantitative assessment using a meta-analysis, given the limited number of studies providing sufficient data for this purpose.

## 5. Discussion

This paper provides a systematic literature review and two meta-analyses on VR applications for safety training in different domains. This was done by analyzing 52 papers published between 2013 and 2021 that met the inclusion criteria and selection process described in Section 3.2. The analysis of these selected papers was guided by the 9 questions

**Table 3**
List of studies comparing different VR training methods (a > b: a performs better than b; a < b: b performs worse than a; = no difference).

| Article | Types of VR | Measurements | Results |
|---|---|---|---|
| (Khan et al., 2021) | Screen-based (multi-screen) vs. screen-based (desktop) | Knowledge acquisition | Desktop > multi-screen |
| (Jung and Ahn, 2018) | Headset-based (HMD) vs. scree-based (desktop) | Knowledge acquisition | Headset-based>/<*screen-based |
| (Maillot et al., 2017) | Projection-based vs. screen-based | Knowledge acquisition | Projection-based > screen-based (for older participants only) |
| (Saeidi et al., 2019) | Headset-based vs. screen-based | User experiences | Headset-based > Screen-based |
| (Grabowski and Jankowski, 2015) | Headset-based (wide FOV) vs. headset-based (narrow FV) | User experiences | High FOV > low FOV |
| (Kim et al., 2020) | Headset-based (augment) vs. headset-based (only) | User experiences | Augment HMD > general HMD |
| (Bhagwat et al., 2021) | Headset-based (game-based scenario) vs. headset-based (VR tour) | User experiences | Game-based >/<* VR tour |
| (Pham et al., 2018a) | Screen-based interactive constructive safety education (eSEC) vs. screen-based (3D VR) | Knowledge acquisition<br>User experiences | eSEC system > 3D VR<br>eSEC system > 3D VR |
| (Dhalmahapatra et al., 2021) | Headset-based (HMD) vs. screen-based (desktop) | Knowledge acquisition<br>User experiences | HMD > desktop<br>= |
| (Buttussi and Chittaro, 2018) | Headset-based (narrow FOV) vs. headset-based (wide FOV) vs. screen-based (desktop) | Knowledge acquisition<br>User experiences | =<br>Wide FOV > narrow FOV |
| (Buttussi and Chittaro, 2021) | Headset-based (HMD) vs. screen-based (smartphone) | Knowledge acquisition<br>User experiences | =<br>HMD > smartphone |
| (Makransky et al., 2019) | Headset-based (HMD) vs. screen-based (desktop) | Knowledge acquisition<br>User experiences | =<br>= |
| (Morélot et al., 2021) | Projection-based vs. screen-based (desktop) | Knowledge acquisition<br>User experiences | Projection-based > screen-based<br>= |
| (Burigat and Chittaro, 2016) | Headset-based (active) vs. headset-based (passive) | Knowledge acquisition<br>User experiences | =<br>Active>/=* passive |
| (Prendinger et al., 2016) | Screen-based (dynamic feedback) vs. screen-based (static feedback); mouse vs. Kinect | Knowledge acquisition<br>User experiences | Dynamic>/=*static<br>Mouse > Kinect |
| (Vahdatikhaki et al., 2019) | Headset-based (context-realistic) vs. headset-based (non-context-realistic) | Knowledge acquisition<br>User experiences | Context-realistic > non-context-realistic<br>Context-realistic>/=* non-context-realistic |

*Various forms of knowledge acquisition and user experiences yielded different outcomes when compared different types of VR, or compared the same device with a different software solution.

listed in Section 3.1. This study provides a more comprehensive analysis of the existing VR applications for safety training compared to existing review studies (Stefan et al., 2023). It provides new insights into how these VR applications were prototyped, considering the hardware solutions and the theories underpinning their development. Further, it provides one of the first meta-analyses providing a quantitative comparison.

This study identified 14 domains and topics where VR safety training has been applied. The primary domains explored in the selected papers are construction and fire safety training, and their prevalence has increased since 2018 (see Section 4.1 and Fig. 2). Construction applications account for 31 % of the examined sample, followed by fire safety applications at 19 %. In 2021, fire and construction training comprised nearly 70 % of all the proposed applications. The third most prominent domain is aviation, representing 10 % of the applications. However, aviation applications are not consistently present throughout the entire investigation period spanning from 2013 to 2021. These results partially align with those reported by Stefan et al. (Stefan et al., 2023), who found that construction applications were among the most popular between 2016 and 2021. Their study also reports that the use of VR in health services is dominant among all domains (36.03 %). However, these VR studies targeted training on healthcare and medical services, not safety training. In our work, we found several healthcare-related VR studies while searching the literature; however, they did not meet the inclusion criteria since they were not about safety training.

Regarding the adoption of theories, only a few studies explicitly demonstrated that the research was supported by a specific theory. Only 5 of the 52 studies (9.6 %) adopted a theory. The theories adopted and identified in these papers are Kirkpatrick's level of training criteria, behavioural skills training, interest theory, self-determination theory, cognitive theory of multimedia learning, cognitive load theory, embodied cognition theory, social cognitive theory, theory of planned behaviour, and game-based learning. Interestingly, many studies did not explicitly report the use of theories; however, they still evaluated their VR safety training applications using the measures and scales that came from different theories. For instance, the measurement of self-efficacy arises from the theory of motivation, the measurement of cognitive load is part of the cognitive load theory, and the investigation on engagement, motivation, and failures can be attributed to the game-based learning theory. As such, this study found a need for future VR safety training studies to focus more on applying theories to shape the design and evaluation of applications. The theories can provide theoretical foundations for designing, developing, or evaluating VR safety training applications. An example of using theories in the design and evaluation can be found in a serious game study for emergency preparedness (Chittaro and Sioni, 2015), where authors used the protection motivation theory to determine the gamification and simulation elements and the measurement of training effectiveness.

Different types of hardware solutions have been used in the selected studies. These can be categorised into headsets, screen-based and projection-based (see Section 4.3). These hardware solutions – from non-immersive to fully immersive systems – can generate different levels of immersion depending on individual factors and system setup (display type and navigation mode, for example) (Paes et al., 2021). Fig. 4 shows that headset setup became the dominant hardware solution in 2019. This trend can be explained by the increment in hardware options and performance and the decrease in prices (Brettschuh et al., 2022). In turn, projection-based setups have not been adopted consistently over the time frame under investigation. The main reason could be related to the costs and space required for this type of system.

Findings show that the proposed applications in the 52 papers were tested in terms of knowledge acquisition and retention as well as user experience. While various solutions are adopted to measure knowledge acquisition and user experience, our findings show that only 36 % of these studies provide retention measurements. The most common retention time in these studies is between 2 and 3 weeks (Fig. 5). This represents a major limitation in the field, as it is essential to investigate the long-term effects of safety training when assessing the effectiveness of a safety training solution. Retention measurements are even more relevant when comparing VR training with traditional training methods. The difference is more evident when focusing on long-term effects, as shown in some of the studies investigated in this review (Chittaro and Buttussi, 2015; Lovreglio et al., 2021).

In this review, the most common learning outcomes delivered by VR safety training (i.e., hazard recognition and management, safety actions, and proper use of equipment and tools) were identified. These findings align with the results by Stefan et al. (Stefan et al., 2023). Further, we also identified that providing immediate feedback for action, regardless of the consequence, was a standard instructional method. This solution was particularly in response to erroneous actions. The effectiveness of using immediate feedback has been investigated by Feng et al., (2023) and Burigat and Chittaro (2016). These studies suggest that providing immediate feedback is an effective way to deliver knowledge and improve self-efficacy in safety training.

Findings show that 15 studies mainly focus on the training prototype challenges instead of comparing VR solutions with traditional methods. This study provides qualitative and quantitative comparisons between VR safety training and traditional safety training. It identified 13 studies comparing these two approaches in terms of knowledge acquisition and only 8 studies in terms of knowledge retention. The meta-analysis results in Section 4.8 highlight that VR solutions performed better than traditional solutions for both knowledge acquisition and retention. However, the analysis highlights a significant level of heterogeneity. This limitation can be overcome in the future using a meta-regression when more studies comparing VR and traditional solutions become available in the literature. However, our findings show similarities with other meta-analyses carried out for other types of training, such as nursing (Chen et al., 2020; Woon et al., 2021), social skill development (Howard and Gutworth, 2020), surgery (Portelli et al., 2020), and education (Kyaw et al., 2019; Merchant et al., 2014; Villena-Taranilla et al., 2022). When comparing the effectiveness of different VR setups, our qualitative analysis in Section 4.9 shows that immersive solutions are more effective. These results are similar to the ones published by Wu et al. (2020).

This review has a few limitations. This review excluded the publications that did not have follow-up experiments or case studies. These studies could eventually provide more detailed insights into the understanding, development, and implementation of VR safety training. Such studies may help address research gaps more comprehensively and gain deeper insights, potentially leading to the generation of additional research questions. Furthermore, medical- and military-related research were excluded from this study because they typically have distinct focuses and very specific professional requirements for their participants. Nonetheless, there could be general safety training knowledge in such studies that could have helped address our research questions and align with our research aims and objectives.

The findings offer potential directions for future studies. Less than half of the studies have conducted retention tests, which were applied within less than four weeks after training. Future studies may consider assessing retention at multiple intervals over longer periods – as observed in longitudinal studies. This approach will provide stronger evidence regarding the reasonable time intervals between training sessions. Additionally, half of the papers come from the construction and fire disciplines. It is worth investigating the effectiveness of VR safety training across other domains with high workplace accident or injury rates.

## 6. Conclusion

In this study, we provide a systematic literature review and two meta-analyses on VR applications for safety training, gathering data from 52 papers published between 2013 and 2021.

This work identified 14 primary domains where VR safety training has been applied, with construction and fire safety training being the most prevalent. Only a small percentage (9.6 %) of the studies explicitly adopted theories, indicating a need for more theoretical grounding in future research. Hardware solutions were crucial factors that have evolved, with headsets becoming the dominant choice in recent years. The most common learning outcomes included hazard recognition, safety actions, and equipment usage, often combined with immediate feedback to enhance knowledge acquisition and self-efficacy. Notably, retention measurements were found to be limited in the reviewed studies, highlighting the need for a more comprehensive investigation of long-term training effects. The meta-analyses revealed that VR safety training outperformed traditional methods in terms of knowledge acquisition and knowledge retention, aligning with previous research findings in different domains. Immersive solutions were shown to be more effective.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

Dr. Lovreglio thanks the NZ Ministry of Business, Innovation & Employment and Royal Society Te Apārangi for funding his research thorough the Marsden Fast Start (MAU2204) and the Rutherford Discovery Fellowship (RDF-MAU2201). The authors thanks Prof. Luca Chittaro (University of Udine) and A/Prof Samart Moodleah (King Mongkut's Institute of Technology Ladkrabang) for helping with the meta analysis sharing their data.

Mrs. Scorgie thanks Cuiying Mu (Library Team - Massey University) for helping her support during the paper search and selection.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssci.2023.106372.

## References

Afzal, M., Shafiq, M.T., 2021. Evaluating 4d-bim and vr for effective safety communication and training: A case study of multilingual construction job-site crew. Buildings 11 (8), 319. https://doi.org/10.3390/buildings11080319.

Ajzen, I., 1991. The theory of planned behavior. Organ. Behavior Hum. Dec. Processes 50 (2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Andrade, C., 2020. Mean difference, standardised mean difference (SMD), and their use in meta-analysis: as simple as it gets. J. Clin. Psychiatry 81 (5), 11349. https://doi.org/10.4088/JCP.20f13681.

Angelov, V., Petkov, E., Shipkovenski, G., Kalushkov, T., 2020. Modern Virtual Reality Headsets. In: 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (Hora 2020), pp. 520-524. https://doi.org/10.1109/HORA49412.2020.9152604.

Asghar, I., Egaji, O.A., Dando, L., Griffiths, M., Jenkins, P., Assoc, M., 2019. Comp A Virtual Reality Based Gas Assessment Application for Training Gas Engineers 2019 Prague, CZECH REPUBLIC. https://doi.org/10.1145/3357419.3357443.

Avveduto, G., Tanca, C., Lorenzini, C., Tecchia, F., Carrozzino, M., Bergamasco, M., 2017. Safety Training Using Virtual Reality: A Comparative Approach. In: De Paolis, L., Bourdot, P., Mongelli, A. (Eds.), Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2017. Lecture Notes in Computer Science, vol 10324. Springer, Cham. https://doi.org/10.1007/978-3-319-60922-5_11.

Bandura, A., Freeman, W.H., Lightsey, R., 1999. Self-efficacy: The exercise of control. J. Cognit. Psychother. 13 (2), 158–166. https://doi.org/10.1891/0889-8391.13.2.158.

Bar-Ilan, J., 2008. Which h-index? - A comparison of WoS, Scopus and Google Scholar. Scientometrics: An International Journal for all Quantitative Aspects of the Science of Science, Communication in Science and Science. Policy 74 (2), 257–271. https://doi.org/10.1007/s11192-008-0216-y.

Becker, L.A., 2000. Effect Size (ES). https://www.uv.es/~friasnav/EffectSizeBecker.pdf (accessed September 12, 2023).

Benitti, F.B.V., 2012. Exploring the educational potential of robotics in schools: A systematic review. Comput. Educ. 58 (3), 978–988. https://doi.org/10.1016/j.compedu.2011.10.006.

Berg, L.P., Vance, J.M., 2017. Industry use of virtual reality in product design and manufacturing: a survey. Virtual Reality 21 (1), 1–17. https://doi.org/10.1007/s10055-016-0293-9.

Bhagwat, K., Kumar, P., Delhi, V.S.K., 2021. Usability of visualisation platform-based safety training and assessment modules for engineering students and construction professionals [Article]. J. Civil Eng. Educ. 147 (2), 34. https://doi.org/10.1061/(ASCE)EI.2643-9115.0000034.

Bhoir, S., Esmaeili, B., 2015. State-of-the-Art Review of Virtual Reality Environment Applications in Construction Safety. In: Aei 2015: Birth and Life of the Integrated Building, pp. 457–468. https://doi.org/10.1061/9780784479070.040.

Borenstein, M., Higgins, J.P.T., Hedges, L.V., Rothstein, H.R., 2017. Basics of meta-analysis: I-2 is not an absolute measure of heterogeneity. Res. Synth. Methods 8 (1), 5–18. https://doi.org/10.1002/jrsm.1230.

Brettschuh, S., Holly, M., Hulla, M., Herstatter, P., Pirker, J., 2022. Virtual Reality in Small and Medium-Sized Enterprises: A Systematic Literature Review. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Christchurch, New Zealand, 2022, pp. 502-507, https://doi.org/10.1109/VRW55335.2022.00112.

British Safety Council, Virtual and augmented reality training. https://www.britsafe.org/training-and-learning/immersive-learning/ (accessed September 12, 23).

Burdea, G.C., Coiffet, P., 2003. Virtual reality technology John Wiley & Sons. ISBN 2003 9780471360896.

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.

Buttussi, F., Chittaro, L., 2018. Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. IEEE Trans. Visualizat. Comput. Graph. 24 (2), 1063–1076. https://doi.org/10.1109/TVCG.2017.2653117. Article: 7817889.

Buttussi, F., Chittaro, L., 2021. A Comparison of Procedural Safety Training in Three Conditions: Virtual Reality Headset, Smartphone, and Printed Materials. IEEE Trans. Learn. Technol. 14 (1), 1–15. Article: 9239901.

Çakiroğlu, Ü., Gökoğlu, S., 2019. Development of fire safety behavioral skills via virtual reality. Comput. Educ. 133, 56–68. https://doi.org/10.1016/j.compedu.2019.01.014.

Chen, F.-Q., Leng, Y.-F., Ge, J.-F., Wang, D.-W., Li, C., Chen, B., Sun, Z.-L., 2020. Effectiveness of virtual reality in nursing education: Meta-analysis. J. Med. Internet Res 22 (9), e18290. https://doi.org/10.2196/18290.

Chen, Y., Wang, X., Xu, H., 2021. Human factors/ergonomics evaluation for virtual reality headsets: a review. CCF Trans. Pervasive Comput. Interaction 3 (2), 99–111. https://doi.org/10.1007/s42486-021-00062-6.

Cherrett, T., Wills, G., Price, J., Maynard, S., Dror, I.E., 2009. Making training more cognitively effective: Making videos interactive. Br. J. Educ. Technol. 40 (6), 1124–1134. https://doi.org/10.1111/j.1467-8535.2009.00985.x.

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.

Chittaro, L., Buttussi, F., Zangrando, N., 2014. Desktop virtual reality for emergency preparedness: User evaluation of an aircraft ditching experience under different fear arousal conditions. In: *ProceedIngs of the ACM Symposium on Virtual Reality Software and Technology*, pp. 141–150. https://doi.org/10.1145/2671015.2671025.

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.

Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., Boyle, J.M., 2012. A systematic literature review of empirical evidence on computer games and serious games. Comput. Educ. 59 (2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004.

Creagh, H., 2003. Cave automatic virtual environment. In: Proceedings: Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Technology Conference, pp. 499-504. https://doi.org/10.1109/EICEMC.2003.1247937.

de Vasconcelos, G.N., Malard, M.L., Van Stralen, M.d.S., Campomori, M.J.L., de Abreu, S. C., Lobosco, T., Gomes, I.F., Lima, L.D.C., 2019. Do we still need CAVEs? In: Architecture in the age of the 4th industrial revolution - in proceedings of the 37th international conference on education and research in computer aided architectural design in Europe.

Dhalmahapatra, K., Maiti, J., Krishna, O.B., 2021. Assessment of virtual reality based safety training simulator for electric overhead crane operations. Saf. Sci. 139 https://doi.org/10.1016/j.ssci.2021.105241.

Durlak, J.A., 2009. How to Select, Calculate, and Interpret Effect Sizes. J. Pediatr. Psychol. 34 (9), 917–928. https://doi.org/10.1093/jpepsy/jsp004.

Dzeng, R.J., Hsueh, H.H., Chang, R.N., 2015. 3D Game-based Training System for Hazard Identification on Construction Site. In: 2015 12th international conference on fuzzy systems and knowledge discovery (fskd)]. https://doi.org/10.1109/FSKD.2015.7382339.

Eiris, R., Gheisari, M., Esmaeili, B., 2020. Desktop-based safety training using 360-degree panorama and static virtual reality techniques: A comparative experimental study. Autom. Constr. 109, 102969 https://doi.org/10.1016/j.autcon.2019.102969.

Feng, Z., González, V.A., Mutch, C., Amor, R., Cabrera-Guerrero, G., 2021. Instructional mechanisms in immersive virtual reality serious games: Earthquake emergency training for children. J. Comput. Assist. Learn. 37 (2), 542–556. https://doi.org/10.1111/jcal.12507.

Feng, Z.A., Gonzalez, V.A., Amor, R., Lovreglio, R., Cabrera-Guerrero, G., 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.

Feng, Z., Gonzalez, A., Amor, R., Soearpoint, 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. Adv. Eng. Inf. 38, 670–682. https://doi.org/10.1016/j.aei.2018.08.018.

Feng, Z., González, V.A., Mutch, C., Amor, R., Cabrera-Guerrero, G., 2023. Exploring spiral narratives with immediate feedback in immersive virtual reality serious games for earthquake emergency training. Multimedia Tools Appl.: Int. J. 82 (1), 125–147. https://doi.org/10.1007/s11042-022-13306-z.

Fire safety in the workplace. https://www.gov.uk/workplace-fire-safety-your-responsibilities/fire-safety-equipment-drills-and-training.

Forst, L., Ahonen, E., Zanoni, J., Holloway-Beth, A., Oschner, M., Kimmel, L., Martino, C., Rodriguez, E., Kader, A., Ringholm, E., Sokas, R., 2013. More than training: Community-based participatory research to reduce injuries among hispanic construction workers. Am. J. Ind. Med. 56 (8), 827–837. https://doi.org/10.1002/ajim.22187.

Gagliardi, E., Bernardini, G., Quagliarini, E., Schumacher, M., Calvaresi, D., 2023. Characterisation and future perspectives of Virtual Reality Evacuation Drills for safe built environments: A Systematic Literature Review. Saf. Sci. 163 https://doi.org/10.1016/j.ssci.2023.106141.

Gao, Y.F., Gonzalez, 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. Comput. Educ. 138, 101–115. https://doi.org/10.1016/j.compedu.2019.05.003.

Grabowski, A., Jankowski, J., 2015. Virtual Reality-based pilot training for underground coal miners [Article]. Saf. Sci. 72, 310–314. https://doi.org/10.1016/j.ssci.2014.09.017.

Granic, I., Lobel, A., Engels, R.C.M.E., 2014. The benefits of playing video games. Am. Psychol. 69 (1), 66–78. https://doi.org/10.1037/a0034857.

Gwynne, S.M.V., Kuligowski, E.D., Boyce, K.E., Nilsson, D., Robbins, A.P., Lovreglio, R., Thomas, J.R., Roy-Poirier, A., 2019. Enhancing egress drills: Preparation and assessment of evacuee performance. Fire Mater. 43 (6), 613–631. https://doi.org/10.1002/fam.2448.

Hale, A.R., 1984. Is safety training worthwhile? J. Occup. Accid. 6, 17–33. https://doi.org/10.1016/0376-6349(84)90026-9.

Harfield, T., Panko, M., Davies, K., Kenley, R., 2007. Toward a learning-styles profile of construction students: Results from New Zealand [Article]. Int. J. Construct. Educ. Res. 3 (3), 143–158.

Havig, P., McIntire, J., Geiselman, E., 2011. Virtual reality in a cave: limitations and the need for HMDs? In: Head-and helmet-mounted displays XVI: design and applications. https://doi.org/10.1117/12.883855.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. Bmj-British Medical Journal 327 (7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Himle, M.B., Miltenberger, R.G., 2004. Preventing unintentional firearm injury in children: The need for behavioral skills training. Educ. Treat. Child. 161–177.

Howard, M.C., Gutworth, M.B., 2020. A meta-analysis of virtual reality training programs for social skill development. Comput. Educ. 144, 103707 https://doi.org/10.1016/j.compedu.2019.103707.

Hsu, E.B., Li, Y., Bayram, J.D., Levinson, D., Yang, S., Monahan, C., 2013. State of virtual reality based disaster preparedness and response training. PLoS currents 5. https://doi.org/10.1371/currents.dis.1ea2b2e71237d5337fa53982a38b2aff.

Ito, K., Tada, M., Ujike, H., Hyodo, K., 2021. Effects of the weight and balance of head-mounted displays on physical load. Appl. Sci. 11 (15), 6802. https://doi.org/10.3390/app11156802.

Jacobsen, E.L., Solberg, A., Golovina, O., Teizer, J., 2021. Active personalised construction safety training using run-time data collection in physical and virtual reality work environments. Constr. Innov. https://doi.org/10.1108/CI-06-2021-0113.

Jeelani, I., Han, K., Albert, A., 2020. Development of virtual reality and stereo-panoramic environments for construction safety training. Eng. Constr. Archit. Manag. 27 (8), 1853–1876. https://doi.org/10.1108/ECAM-07-2019-0391.

Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A., 2008. Measuring and defining the experience of immersion in games. Int. J. Hum Comput Stud. 66 (9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004.

Jeon, J.C., 2016. Development and effect analysis of experiential electrical safety education system based on virtual reality [Article]. Trans. Kor. Inst. Elect. Eng. 65 (10), 1767–1773. https://doi.org/10.5370/KIEE.2016.65.10.1767.

Jung, J., Ahn, Y.J., 2018. Effects of interface on procedural skill transfer in virtual training: Lifeboat launching operation study. Comput. Anim. Virtual Worlds 29 (3/4), 1-N.PAG. https://doi.org/10.1002/cav.1812.

Kanade, S.G., Duffy, V.G., 2022. Use of Virtual Reality for Safety Training: A Systematic Review. Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Health, Operations Management, and Design, Pt Ii, 13320, 364-375. https://doi.org/10.1007/978-3-031-06018-2_25.

Khan, K.S., Kunz, R., Kleijnen, J., Antes, G., 2003. Five steps to conducting a systematic review. J. R. Soc. Med. 96 (3), 118–121. https://doi.org/10.1177/014107680309600304.

Khan, N., Muhammad, K., Hussain, T., Nasir, M., Munsif, M., Imran, A.S., Sajjad, M., 2021. An adaptive game-based learning strategy for children road safety education and practice in virtual space. Sensors 21 (11), 3661. https://doi.org/10.3390/s21113661.

Kim, J., Ahn, S., Han, S., 2020. Impact of VR Settings on Immersiveness in VR-Based Construction Skills Training: Case Study of Scaffolding Work. In: Construction Research Congress 2020: Computer Applications, pp. 336-344. https://doi.org/10.1061/9780784482865.036.

Kinateder, M., Pauli, P., Müller, M., Krieger, J., Heimbecher, F., Rönnau, I., Bergerhausen, U., Vollmann, G., Vogt, P., Mühlberger, A., 2013. Human behaviour in severe tunnel accidents: Effects of information and behavioural training. Transp. Res. F Psychol. Behav. 17, 20–32. https://doi.org/10.1016/j.trf.2012.09.001.

Kirkpatrick, J., 2015. An introduction to the new world Kirkpatrick model. Kirkpatrick Partners, 2019. http://www.kirkpatrickpartners.com/wp-content/uploads/2021/11/Introduction-to-the-Kirkpatrick-New-World-Model.pdf (accessed September 12, 2023).

Kyaw, B.M., Saxena, N., Posadzki, P., Vseteckova, J., Nikolaou, C.K., George, P.P., Divakar, U., Masiello, I., Kononowicz, A.A., Zary, N., 2019. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. J. Med. Internet Res. 21 (1), e12959.

Lacko, J., 2020. Health safety training for industry in virtual reality. In: Proceedings of the 30th International Conference on Cybernetics and Informatics 2020, Velke Karlovice, Czech Republic, 2020, pp. 1-5, https://doi.org/10.1109/KI48306.2020.9039854.

LaValle, S.M., 2023. Virtual reality. Cambridge university press. ISBN: 1107198933.

Lee, K.-W., Lee, D.-S., 2004. Development of a CAVE type Virtual Reality System for 3-D Spatial Data Visualization. Geophys. Geophys. Explor. 7 (2), 117–120.

Li, C., Liang, W., Quigley, C., Zhao, Y., Yu, L.F., 2017. Earthquake Safety Training through Virtual Drills. IEEE Trans. Vis. Comput. Graph. 23 (4), 1275–1284. https://doi.org/10.1109/TVCG.2017.2656958. Article 7829397.

Li, X., Yi, W., Chi, H.L., Wang, X.Y., Chan, A.P.C., 2018. A critical review of virtual and augmented reality (VR/AR) applications in construction safety. Autom. Constr. 86, 150–162. https://doi.org/10.1016/j.autcon.2017.11.003.

Liang, Z., Zhou, K., Gao, K., 2019. Development of Virtual Reality Serious Game for Underground Rock-Related Hazards Safety Training. IEEE Access, 7, 118639-118649, Article 8795446. https://doi.org/10.1109/ACCESS.2019.2934990.

Liang, H., Ge, C., Liang, F., Sun, Y., 2020. VR-based training model for enhancing fire evacuee safety. Int. J. Performab. Eng., 16(1), 107-117-117. https://doi.org/10.23940/ijpe.20.01.p12.107117.

Loosemore, M., Malouf, N., 2019. Safety training and positive safety attitude formation in the Australian construction industry. Saf. Sci. 113, 233–243. https://doi.org/10.1016/j.ssci.2018.11.029.

Lovreglio, R., Duan, X., Rahouti, A., Phipps, R., Nilsson, D., 2021. Comparing the effectiveness of fire extinguisher virtual reality and video training. Virtual Reality 25 (1), 133–145. https://doi.org/10.1007/s10055-020-00447-5.

Lovreglio, R., 2020. Virtual and Augmented reality for human behaviour in disasters: a review. In: Fire and Evacuation Modeling Technical Conference (FEMTC) 2020 Online Conference.

Lu, X., Davis, S., 2018. Priming effects on safety decisions in a virtual construction simulator. Eng. Constr. Archit. Manag. 25 (2), 273–294. https://doi.org/10.1108/ECAM-05-2016-0114.

Maillot, P., Dommes, A., Dang, N.T., Vienne, F., 2017. Training the elderly in pedestrian safety: Transfer effect between two virtual reality simulation devices. Accid. Anal. Prev. 99, 161–170. https://doi.org/10.1016/j.aap.2016.11.017.

Makransky, G., Borre-Gude, S., Mayer, R.E., 2019. Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. J. Comput. Assist. Learn. 35 (6), 691–707. https://doi.org/10.1111/jcal.12375.

Mayer, R.E., 2005. Cognitive theory of multimedia learning. The Cambridge handbook of multimedia learning, 41, 31-48. ISBN: 1139992481.

Menzemer, L.W., Ronchi, E., Karsten, M.M.V., Gwynne, S., Frederiksen, J., 2023. A scoping review and bibliometric analysis of methods for fire evacuation training in buildings. Fire Saf. J., 103742

Merchant, Z., Goetz, E.T., Cifuentes, L., Keeney-Kennicutt, W., Davis, T.J., 2014. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. Comput. Educ. 70, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033.

Morélot, S., Garrigou, A., Dedieu, J., N'Kaoua, B., 2021. Virtual reality for fire safety training: Influence of immersion and sense of presence on conceptual and procedural acquisition. Comput. Educ. 166 https://doi.org/10.1016/j.compedu.2021.104145.

Nazli, N.N.N.N., Sipon, S., Radzi, H.M., 2014. Analysis of Training Needs in Disaster Preparedness. Procedia. Soc. Behav. Sci. 140, 576–580. https://doi.org/10.1016/j.sbspro.2014.04.473.

Nykänen, M., Puro, V., Tiikkaja, M., Kannisto, H., Lantto, E., Simpura, F., Uusitalo, J., Lukander, K., Räsänen, T., Heikkilä, T., Teperi, A.M., 2020. Implementing and evaluating novel safety training methods for construction sector workers: Results of a randomised controlled trial. J. Saf. Res. 75, 205–221. https://doi.org/10.1016/j.jsr.2020.09.015.

Oliva, D., Somerkoski, B., Tarkkanen, K., Lehto, A., Luimula, M., 2019. Virtual reality as a communication tool for fire safety - Experiences from the VirPa project GamiFIN Conference 2019 Levi, Finland.

Paes, D., Irizarry, J., Pujoni, D., 2021. An evidence of cognitive benefits from immersive design review: Comparing three-dimensional perception and presence between immersive and non-immersive virtual environments. Autom. Constr. 130 https://doi.org/10.1016/j.autcon.2021.103849.

Parliamentary Counsel's Committee, 2023. Model Work Health and Safety Regulations. https://www.safeworkaustralia.gov.au/sites/default/files/2023-08/model-whs-regulations-1_august_2023.pdf.

Paszkiewicz, A., Salach, M., Dymora, P., Bolanowski, M., Budzik, G., Kubiak, P., 2021. Methodology of implementing virtual reality in education for industry 4.0. Sustainability (Switzerland) 13 (9), 5049. https://doi.org/10.3390/su13095049.

Pedram, S., Perez, P., Palmisano, S., Farrelly, M., 2017. Evaluating 360-Virtual Reality for Mining Industry's Safety Training. In: Stephanidis, C. (eds) HCI International 2017 – Posters' Extended Abstracts. HCI 2017. Communications in Computer and Information Science, vol 713. Springer, Cham. https://doi.org/10.1007/978-3-319-58750-9_77.

Pedram, S., Palmisano, S., Skarbez, R., Perez, P., Farrelly, M., 2020. Investigating the process of mine rescuers' safety training with immersive virtual reality: A structural equation modelling approach. Comput. Educ. 153, 103891 https://doi.org/10.1016/j.compedu.2020.103891.

Pham, H.C., Dao, N.N., Kim, J.U., Cho, S., Park, C.S., 2018a. Energy-efficient learning system using Web-based panoramic virtual photoreality for interactive construction safety education. Sustainability (Switzerland) 10 (7), 2262. https://doi.org/10.3390/su10072262.

Pham, H.C., Dao, N.N., Pedro, A., Le, Q.T., Hussain, R., Cho, S., Park, C.S., 2018b. Virtual field trip for mobile construction safety education using 360-degree panoramic virtual reality. Int. J. Eng. Educ. 34 (4), 1174–1191. https://www.researchgate.net/profile/Nhu-Ngoc-Dao/publication/326096578_Virtual_Field_Trip_for_Mobile_Construction_Safety_Education_Using_360-Degree_Panoramic_Virtual_Reality/links/5b4413f30f7e9bb59b1b1537/Virtual-Field-Trip-for-Mobile-Construction-Safety-Education-Using-360-Degree-Panoramic-Virtual-Reality.pdf.

Plass, J.L., Homer, B.D., Kinzer, C.K., 2015. Foundations of game-based learning. *Educational psychologist* 50 (4), 258–283. https://doi.org/10.1080/00461520.2015.1122533.

Portelli, M., Bianco, S., Bezzina, T., Abela, J., 2020. Virtual reality training compared with apprenticeship training in laparoscopic surgery: a meta-analysis. Ann. Roy. College Surgeons England 102 (9), 672–684. https://doi.org/10.1308/rcsann.2020.0178.

Poyade, M., Eaglesham, C., Trench, J., Reid, M., 2021. A Transferable Psychological Evaluation of Virtual Reality Applied to Safety Training in Chemical Manufacturing. J. Chem. Health Saf. 28 (1), 55–65. https://doi.org/10.1021/acs.chas.0c00105.

Prendinger, H., Alvarez, N., Sanchez-Ruiz, A., Cavazza, M., Catarino, J., Oliveira, J., Prada, R., Fujimoto, S., Shigematsu, M., 2016. Intelligent biohazard training based on real-time task recognition. ACM Trans. Interact. Intell. Syst. 6 (3), Article 21. https://doi.org/10.1145/2883617.

Rahmalan, H., Mohamad, S.N.M., Rahman, A.F.N.A., Aziz, A., Ganasan, A., 2020. Development of virtual reality training for fire safety education. Int. J. Adv. Trends Comput. Sci. Eng., 9(4), 5906-5912, Article 253. https://doi.org/10.30534/ijatcse/2020/253942020.

Rahouti, A., Lovreglio, R., Datoussaïd, S., Descamps, T., 2021. Prototyping and Validating a Non-immersive Virtual Reality Serious Game for Healthcare Fire Safety Training. Fire Technol. 1–38 https://doi.org/10.1007/s10694-021-01098-x.

Rendevski, N., Trajcevska, D., Dimovski, M., Veljanovski, K., Popov, A., Emini, N., Veljanovski, D., 2022. PC VR vs Standalone VR Fully-Immersive Applications: History, Technical Aspects and Performance. In: 2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST).

Renninger, K.A., Hidi, S., 2015. The power of interest for motivation and engagement. Routledge.

Robertson, G.G., Card, S.K., Mackinlay, J.D., 1993. Three views of virtual reality: nonimmersive virtual reality. Computer 26 (2), 81.

Robson, L.S., Stephenson, C.M., Schulte, P.A., Amick, B.C., Irvin, E.L., Eggerth, D.E., Chan, S., Bielecky, A.R., Wang, A.M., Heidotting, T.L., Peters, R.H., Clarke, J.A., Cullen, K., Rotunda, C.J., Grubb, P.L., 2012. A systematic review of the effectiveness of occupational health and safety training. Scand. J. Work Environ. Health 38 (3), 193–208. https://doi.org/10.5271/sjweh.3259.

Rosenthal, L., Schisterman, E., 2010. Meta-analysis: drawing conclusions when study results vary. Adv. Protocols Oxidat. Stress II, 427–434.

Ryan, R.M., Deci, E.L., 2016. Facilitating and hindering motivation, learning, and well-being in schools: Research and observations from self-determination theory. In: Handbook of motivation at school. Routledge, pp. 96–119.

Sacks, R., Perlman, A., Barak, R., 2013. Construction safety training using immersive virtual reality. Constr. Manag. Econ. 31 (9), 1005–1017. https://doi.org/10.1080/01446193.2013.828844.

Saeidi, S., Zhu, Y., Lifkooee, M.Z., Mollazadeh, M., Li, X., 2019. Co-Presence in a Shared Virtual Environment (SVE): A Case Study of Highway Work Zone Construction. In: ICCREM 2019: Innovative Construction Project Management and Construction Industrialisation - Proceedings of the International Conference on Construction and Real Estate Management 2019.

Saghafian, M., Laumann, K., Akhtar, R.S., Skogstad, M.R., 2020. The Evaluation of Virtual Reality Fire Extinguisher Training. Front. Psychol. 11, 593466 https://doi.org/10.3389/fpsyg.2020.593466.

Sandin, D.J., DeFanti, T.A., Cruz-Neira, C., 1993. Surround-screen projection-based virtual reality: The design and implementation of the CAVE.

Satapanasatien, K., Phuawiriyakul, T., Moodleah, S., 2021. A Development of Game-Based Learning in Virtual Reality for Fire Safety Training in Thailand. In: JCSSE 2021 - 18th International Joint Conference on Computer Science and Software Engineering, Lampang, Thailand, 2021, pp. 1-6. https://doi.org/10.1109/JCSSE53117.2021.9493836.

Sayli, S., Luis, R., Fahad, A., Khalid, N., 2021. Virtual World as an Interactive Safety Training Platform. Information 12 (219), 219. https://doi.org/10.3390/info12060219.

Schwebel, D.C., Combs, T., Rodriguez, D., Severson, J., Sisiopiku, V., 2016. Community-based pedestrian safety training in virtual reality: A pragmatic trial. Accid. Anal. Prev. 86, 9–15. https://doi.org/10.1016/j.aap.2015.10.002.

Schwebel, D.C., Wu, Y., Li, P., Severson, J., He, Y.F., Xiang, H., Hu, G.Q., 2018. Featured Article: Evaluating Smartphone-Based Virtual Reality to Improve Chinese Schoolchildren's Pedestrian Safety: A Nonrandomised Trial [Article]. J. Pediatr. Psychol. 43 (5), 473–484. https://doi.org/10.1093/jpepsy/jsx147.

Shadish, W.R., Hedges, L.V., Pustejovsky, J.E., 2014. Analysis and meta-analysis of single-case designs with a standardised mean difference statistic: A primer and applications. J. Sch. Psychol. 52 (2), 123–147. https://doi.org/10.1016/j.jsp.2013.11.005.

Shiradkar, S., Rabelo, L., Alasim, F., Nagadi, K., 2021. Virtual world as an interactive safety training platform. Information (Switzerland) 12 (6), 219. https://doi.org/10.3390/info12060219.

Site Safe Passport Courses - Third Edition, 2019. https://www.sitesafe.org.nz/globalassets/training/passport-courses-catalogue—mar-20202.pdf.

Stefan, H., Mortimer, M., Horan, B., 2023. Evaluating the effectiveness of virtual reality for safety-relevant training: a systematic review. Virtual Reality 1–31. https://doi.org/10.1007/s10055-023-00843-7.

Stewart, K.K., Carr, J.E., LeBlanc, L.A., 2007. Evaluation of family-implemented behavioral skills training for teaching social skills to a child with Asperger's disorder. Clin. Case Stud. 6 (3), 252–262. https://doi.org/10.1177/1534650106286940.

Sweller, J., 2011. Cognitive load theory. In: Psychology of learning and motivation, vol. 55, pp. 37-76. https://doi.org/10.1016/B978-0-12-387691-1.00002-8.

Takeshima, N., Sozu, T., Tajika, A., Ogawa, Y., Hayasaka, Y., Furukawa, T.A., 2014. Which is more generalisable, powerful and interpretable in meta-analyses, mean difference or standardised mean difference? Bmc Med. Res. Methodol. 14 https://doi.org/10.1186/1471-2288-14-30.

Thornton, A., Lee, P., 2000. Publication bias in meta-analysis: its causes and consequences. J. Clin. Epidemiol. 53 (2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4.

Tichon, J., Burgess-Limerick, R., 2011. A review of virtual reality as a medium for safety related training in mining. J. Health Safety Res. Practice 3 (1), 33–40.

Vahdatikhaki, F., El Ammari, K., Langroodi, A.K., Miller, S., Hammad, A., Doree, A., 2019. Beyond data visualisation: A context-realistic construction equipment training simulators. Autom. Constr. 106, 102853 https://doi.org/10.1016/j.autcon.2019.102853.

Valentine, A., Van der Veen, T., Kenworthy, P., Hassan, G.M., Guzzomi, A., Khan, R.N., Male, S.A., 2021. Using head mounted display virtual reality simulations in large engineering classes: Operating vs observing. Australas. J. Educ. Technol. 37 (3), 119–136. https://doi.org/10.14742/ajet.5487.

Validakis, V., 2015. Virtual reality: The new mine safety training tool. Australian Mining, 107(7), 12-15-15. https://ezproxy.massey.ac.nz/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=edselc&AN=edselc.2-52.0-84953210289&site=eds-live&scope=site&authtype=sso&custid=s3027306.

Villena-Taranilla, R., Tirado-Olivares, S., Cozar-Gutierrez, R., González-Calero, J.A., 2022. Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis. Educ. Res. Rev. 35, 100434 https://doi.org/10.1016/j.edurev.2022.100434.

Wang, Z., Yang, Q., 2019. Scene Building of Fully Mechanised Mining and Safety Training in Dongbaowei Coal Mine Based on VR Technology. In: Proceedings - 2019 International Conference on Artificial Intelligence and Advanced Manufacturing, AIAM 2019, Dublin, Ireland, 349-354. https://doi.org/10.1109/AIAM48774.2019.00075.

Weiss, M.J., 2022. Occupational Safety and Health Act of 1970. In: Salem Press. https://www.salempress.com/Media/SalemPress/samples/american_business_pgs.pdf (accessed September 12, 23).

Wilson, M., 2002. Six views of embodied cognition. Psychon. Bull. Rev. 9, 625–636. https://doi.org/10.3758/BF03196322.

Wilson, C., 2008. Avatars, Virtual Reality Technology, and the U.S. Military: Emerging Policy Issues. Washington, DC: Congressional Research Service, 2008. https://apps.dtic.mil/sti/pdfs/ADA480182.pdf (accessed September 12, 23).

Woon, A.P.N., Mok, W.Q., Chieng, Y.J.S., Zhang, H.M., Ramos, P., Mustadi, H.B., Lau, Y., 2021. Effectiveness of virtual reality training in improving knowledge among nursing students: A systematic review, meta-analysis and meta-regression. Nurse Educ. Today 98, 104655. https://doi.org/10.1016/j.nedt.2020.104655.

Wu, B., Yu, X., Gu, X., 2020. Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. Br. J. Educ. Technol. 51 (6), 1991–2005. https://doi.org/10.1111/bjet.13023.

Wuchty, S., Jones, B.F., Uzzi, B., 2007. The Increasing Dominance of Teams in Production of Knowledge. Science 316 (5827), 1036–1039. https://doi.org/10.1126/science.1136099.

Xueqing, L., Steven, D., 2018. Priming effects on safety decisions in a virtual construction simulator. Eng. Constr. Archit. Manag. 25 (2), 273–294. https://doi.org/10.1108/ECAM-05-2016-0114.

Zhao, D., Lucas, J., 2015. Virtual reality simulation for construction safety promotion. Int. J. Inj. Contr. Saf. Promot. 22 (1), 57–67. https://doi.org/10.1080/17457300.2013.861853.