Human Factors Research in Immersive Virtual Reality Firefighter Training: A Systematic Review

Steven G Wheeler¹, Hendrik Engelbrecht¹ and Simon Hoermann¹,²*

¹HIT Lab NZ, University of Canterbury, Christchurch, New Zealand, ²School of Product Design, University of Canterbury, Christchurch, New Zealand

Immersive virtual reality (VR) shows a lot of potential for the training of professionals in the emergency response domain. Firefighters occupy a unique position among emergency personnel as the threats they encounter are mainly environmental. Immersive VR therefore represents a great opportunity to be utilized for firefighter training. This systematic review summarizes the existing literature of VR firefighting training that has a specific focus on human factors and learning outcomes, as opposed to literature that solely covers the system, or simulation, with little consideration given to its user. An extensive literature search followed by rigorous filtering of publications with narrowly defined criteria was performed to aggregate results from methodologically sound user studies. The included studies provide evidence that suggests the suitability of VR firefighter training, especially in search and rescue and commander training scenarios. Although the overall number of publications is small, the viability of VR as an ecologically valid analog to real-life training is promising. In the future, more work is needed to establish clear evidence and guidelines to optimize the effectiveness of VR training and to increase reliable data through appropriate research endeavors.

Keywords: training, firefighter, human factor, human computer interaction (HCI), virtual reality

1 INTRODUCTION

Virtual reality (VR) technology has been evolving rapidly over the past few years. VR is making its way into the consumer market with affordable headsets in a variety of price ranges and research in the domain of the application of VR is at a record pace (Anthes et al., 2016).

Previous studies suggest that VR is a valuable training tool in the medical, educational, and manufacturing domains, such as the training of laparoscopic surgery (Alaker et al., 2016), in cognitive behavior therapy (Lindner, 2020), the creation of empathy in the user (Kilteni et al., 2012; Shin, 2018), or as a teaching tool in the manufacturing domain (Mujber et al., 2004). Research in the field of military applications has used VR successfully for the treatment of adverse mental conditions (Rizzo et al., 2011) as well as increasing mental preparedness of soldiers (Wiederhold and Wiederhold, 2004; Stetz et al., 2007) (known as stress inoculation training). VR has also been successfully used to teach correct safety procedure in hazardous situations (Ha et al., 2016; Oliva et al., 2019; Ooi et al., 2019).

VR enables users to be placed into a believable, customizable, and controllable virtual environment. Due to this, there is great interest in the educational domain thanks to the possibility of virtual worlds enabling experiential learning. As defined by Kolb (1984), experiential learning is achieved through the transformation of experience into knowledge.
There has been considerable interest in applying virtual worlds for experiential learning; see, for example, Jarmen et al. (2009) or Le et al. (2015).

Applying this to the firefighting context, the possibility of enabling experiential learning in a virtual space is a great opportunity for hands-on training that does not need to be reliant on the personnel, resources, and budget for training firefighters. VR might therefore enable cost-effective and frequent training for a large variety of scenarios. Due to its immersive properties, VR is gaining traction in the training of high-risk job domains. Stimulating the feeling of presence, virtual environments can arouse physiological responses as indicators of stress on par with real-life arousal (Wiederhold et al., 2001; Meehan et al., 2003), which shows promise for VR possibly being an ecologically valid analog to real-life training exercises. Firefighter trainees are faced with a multitude of environmental hazards making the use of VR for training a natural extension of what has been shown in other domains. Yet, with the variety of threats faced, the difference in skills needed and the mental demands seemingly unique, the effectiveness of VR training for firefighting needs to be looked at as an independent investigation.

This article explores and analyzes the field of firefighter VR training using a systematic search procedure. To obtain relevant research that enriches the pool of evidence in this domain, the researchers are purposefully restricting the analysis to research pertaining to the domain of human factors with the goal of assessing the impact on end-users within the target population.

2 DEFINITIONS

2.1 Immersive and Non-Immersive Virtual Reality

For this article, the definition for immersive VR concerns itself with the direct manipulation of the environment using input and visualization devices that respond to the natural motion of the user (Robertson et al., 1993). Several researchers have shown that non-immersive, monitor-bound simulations offer possibilities for training firefighters [see, for example, (St Julien and Shaw, 2003; Yuan et al., 2007; van Berlo et al., 2005)]. However, as immersive VR technology has many distinctive properties and brings with it many unique challenges and considerations—for example, the issue of cybersickness (LaViola, 2000) or the challenge of creating effective input methods in VR (Choe et al., 2019)—we argue that it needs to be treated as a separate inquiry. Therefore, VR setups utilizing head-mounted displays and CAVE systems (Cruz-Neira et al., 1993) are the focus of this inquiry, and desktop monitor-bound simulations are not within the scope of this investigation.

2.2 Presence

Presence is the result of immersion in the virtual environment where the user feels a sense of being physically part of the virtual world as if they have been transported to another place, independent from the current real-world location (Slater and Usoh, 1993; Lombard and Ditton, 1997). Due to this, VR has been shown to be able to stimulate similar responses and behavior in reactions to hazards and risks as they would in real-life (Alcañiz et al., 2009). As such, effective transmission of presence has been found to make VR a safe and effective medium to train personnel in high-risk situations (Amokrane et al., 2008) and, therefore, is an important factor to consider in the discussion of firefighting training—a job domain with a high level of risk to the personnel.

2.3 Ecological Validity

Differing from both immersion and presence, we judge ecological validity to refer to how representative the virtual activities are of their real-life counterparts (Paljic, 2017). As the main focus of this inquiry is specifically looking at VR as a predictive tool for training, we deem it important to consider the ecological validity of each study to judge its efficacy in real-world applications. This is not to be confused with simply considering the physical fidelity, or graphical realism, of the virtual environment, which has been shown to have a limited impact on the user experience (Lukosch et al., 2019). Rather, this article directly considers the input methods used, the equivalent real-world equipment and the relevance of the virtual task to real-world situations.

2.4 Training, Aids, and Post-Hoc Applications

This article looks into the application of training, i.e., the acquisition of mental and physical skills, prior to the usage of such skills in the real world. This means that applications only for the use during deployment are not part of the inquiry, since this review is strictly on the potential for acquisition and training of skills and not the improvement of the execution with the usage of VR technology. The same principles apply to post-hoc applications, which concern themselves with either the treatment or post-incident analysis of factors resulting from the work itself. While there is an overlap between post-hoc applications used to reinforce skills that have already been executed and trained, the focus of these applications is not on the acquisition and maintenance of skills through VR, but represents a combination of approaches. We argue that this, while naturally a part of future inquiries, introduces too much noise into the validation of training in this domain.

2.5 Human Factors Evaluations

In this systematic review, the term “human factors” is being used in relation to the evaluation of behavioral and psychological outcomes of training applications. The term thereby extends functionality considerations beyond a mere systems perspective; the literature that only focuses on the purely functional aspects of training execution in the virtual environment, without considering the end-user, is excluded from this investigation. We aim to clarify this due to some work conflating functionality evaluations with training effectiveness. In these cases, the effect of virtual training execution on the user is often not specifically considered. The successful completion of a virtual task alone is often deemed as proof to the ecological validity of simulation. The impact of integrating existing training routines into virtual worlds needs a
holistic investigation that encompasses functional, as well as psychological and behavioral outcomes for assessing their effectiveness in the human factors domain.

3 POPULATION CONSIDERATIONS

3.1 Emergency Response and VR Research

There has been a lot of interest in VR technology for the training of emergency response employees. For example, the development of VR disaster response scenarios has gained popularity [see, for example, (Chow et al., 2005; Vincent et al., 2008; Sharma et al., 2014)] since it enables cost-effective training of large-scale exercises and offers immersive properties that are difficult to replicate in desktop monitor-bound training.

The term emergency response is an umbrella term that describes any profession that works in the service of public safety and health often under adverse or threatening conditions. Included under this umbrella term are professions such as emergency medical technicians, police officers, or firefighters. While these are all distinct professions, there is an overlap in the kind of situations all three encounter, such as traffic accidents or natural disasters. Hence, research in this domain is often grouped under this umbrella term, with generalizations being made across the entire domain.

While there is an overlap in skills and mental demands, the findings in one area should not be generalized with undue haste to other areas. Emergency medical technicians (EMTs) are primarily faced with mental strains in the form of potentially traumatizing imagery (e.g., in the form of heavily injured patients) at the scene. While there can be threats to EMTs during deployment, sprains and strains are most common and injury rates are potentially lower than those of other emergency response occupations (Heick et al., 2009). The skills needed are largely independent of the environment, as they apply to the handling of the patient directly.

Police officers, on the other hand, often deal with very direct threats in the form of human contact. Suspects, or generally people causing a disturbance, can pose a threat to the officer if the situation gets out of control. The environmental threats faced only account for a small fraction in the case of, for example, traffic accidents or disaster response, with the risk of injury being highest for assaults from non-compliant offenders (Lyons et al., 2017). Similarly to EMT’s, the skills needed are not completely independent of the environment, but interpersonal contact plays the main factor in the everyday life of the police officer when it comes to occupational threats.

This review concerns itself with the application of VR training for firefighters exclusively. The work environment of firefighters is hypothesized to be unique due to the nature of the threats and the skills applied being heavily dependent on the interaction with the environment. Firefighters work in an environment full of dangers. Fire, falling objects, explosions, smoke, and intense heat are only some of the large variety of environmental threats faced (Dunn, 2015). In 2017 alone, a total of 50,455 firefighters were injured during deployment in the United States. Furthermore deployment resulted in 60 deaths in 2017. Even during training itself, 8,380 injuries and ten deaths were recorded in 2017 (Evarts and Molis, 2018; Fathy et al., 2018). Numerous threats are faced by firefighters, and with high potential risk to life and well-being, ecologically valid training is necessary. Training in an environment that adequately represents environmental threats faced during deployment is vital to learning skills.

While a transfer of knowledge gained in any emergency response research can be valuable for informing system design in other areas, the independent aggregation of results remains important for obtaining evidence that can be used as a building block for future work. A high level of scrutiny is required when it comes to the development of new technologies, since the failure to do so can impact the safety of the workforce in the respective occupation. We therefore argue that VR research should treat these occupations as separate fields of inquiry when assessing the impact on human factors.

4 SEARCH STRATEGY

This section describes the details of the publication search and selection strategy and explains the reasons for their application in this systematic review.

4.1 Search-Terms and Databases

Firefighter research within human–computer interaction (HCI) is a multidisciplinary field; hence, this review aims to capture work published in engineering and computer science, as well as in all life-, health-, physical-, and social-sciences fields. While this has resulted in only a few unique additions to the search results, this inclusive approach was chosen to prevent the omission of potentially relevant work. The following databases were used for the systematic search:

- Scopus (Elsevier Publishers, Amsterdam, Netherlands)
- EI Compendex (Elsevier Publishers, Amsterdam, Netherlands)¹
- IEEE Xplore (IEEE, Piscataway, New Jersey, United States)
- PsycINFO (APA, Washington, Washington DC, United States)

For the purpose of this review, we aimed to purposefully narrow the scope of the assessed literature to human factors evaluation of training systems for fire service employees using immersive virtual reality technology. As such, the search terms had to be specified and justified with regard to that goal.

4.1.1 Technology

The value of immersive VR for training simulations lies in the match of immersive properties with the threats faced by the target population. With a large part of the most dangerous threats encountered by firefighters being environmental in nature, there is an opportunity for immersive VR to make a unique contribution to training routines. While mixed reality systems

¹Includes ACM Publications such as proceeding of the Conference on Human Factors in Computing Systems.
might arguably be able to present threats to trainees with similarly high physical fidelity, results obtained from evaluations deploying these technologies in the firefighting domain might not be transferable to immersive VR training and further increase noise for establishing a clear baseline for the utility of this technology.

For this review, the following terms were used as part of the systematic search:

**virtual reality; VR**

4.1.2 Target Population
As discussed previously, the population of firefighters occupies a unique position within the emergency response domain with regard to threats faced and skills needed. To capture the entirety of the target population, the terms used in the search were kept broad and only included a few specialized terms, such as land search and rescue (LandSAR), which revealed additional citations that were not covered by the other, more general, search terms. The broadness of the terms used means that more additional manual processing and filtering of the resulting citations will be needed, but this was deemed necessary to prevent any possible omission of work in this domain.

For this review, the following terms were used as part of the systematic search:

**firefight*; fire service*; fire fight*; fire department; landsar; usar**

4.1.3 Aim
The aim of this article was to capture any possible application of immersive VR systems for training purposes. Training in this case is defined as any form of process applied with the aim of improving skills (mental and physical) or knowledge before they are needed. During preliminary searches, we found that several terms overlapped with the terms already being used, resulting in no new unique citations, and were therefore excluded from the systematic search, namely, teach*, coach*, and instruct*.

For this article, the following terms were used as part of the systematic search:

**train*; educat*; learn*; habituat*; condition*; expos*; treat**

4.2 Selection Criteria

4.2.1 Target Population
The target population of the citation needs to be concerned with fire service employees. This does include any kind of specialization that can be obtained within the fire service and extends throughout ranks. We excluded articles that exclusively investigated other emergency response personnel or unrelated occupations.

4.2.2 Technology Used
Immersive virtual reality, i.e., a CAVE system or head-mounted display, needs to be used as the main technology in the article.
Augmented- or mixed-reality, as well as monitor-bound simulations, are not within the scope of this review.

4.2.3 Practical Application
The aim of this investigation is to evaluate the scope of research done in the domain of human factors research. For an article to be included in this review, it needs to be aimed towards a practical application of technology for the fire service. Pure system articles, e.g., development of algorithms, will be excluded.

4.2.4 Sample
The sample used during evaluation needs to represent the population of firefighters. This does include the approximation of the target population by using civilian participants to act as firefighters. When proxies were used instead of firefighters, this limitation needed to be clearly acknowledged as a potential limitation.

4.2.5 Aim
The research needs to be on a training system that is concerned with the acquisition or maintenance of skills or knowledge before an event demands them during real deployment. Systems intended for use during deployment, e.g., technology to improve operations in real life, or post deployment, e.g., for the treatment of conditions such as PTSD, will be excluded.

4.2.6 Measures
The research needs to evaluate the impact of the system with relevant outcome measures for the human factors domain. Articles with a sole focus on system measures with no, or vastly inadequate, user studies will be excluded from the review.

4.3 Process and Results
The process of the systematic search can be seen in Figure 1.

First, the search terms were defined to specify the scope of the review, while retaining a broad enough search to obtain all relevant literature. Databases were selected based on their coverage of relevant fields with expected redundancy among the results. The search procedure for all databases was kept as similar as possible. The search terms were used to look for matches in the title, abstract or associated keywords of the articles. Only English language documents were included in the review, and appropriate filters were set for all database search engines. While the exact settings differed slightly depending on the database, as certain document types were grouped together, only journal articles, conference articles and review articles published up to the writing of this article were included as part of the review. The total amount of citations identified was 300. After the removal of duplicates, the citation pool was reduced to 168 articles.

Next, for the first round of applying the exclusion criteria, as specified above, the abstracts and conclusions were evaluated and articles were removed accordingly. Afterward, the remaining 110 articles were evaluated based on the full text. Any deviation from the above mentioned criteria resulted in the exclusion of the publication. This was also applicable to work that, for example, failed to describe the demographics of participants entirely (i.e., it is unclear whether members of the target population were sampled) or did not describe what hardware was used for training. The latter becomes especially troubling with the term virtual reality having been used interchangeably with monitor-bound simulations in many bodies of work. In these cases, some articles needed to be excluded, because no further information was given as to whether or not immersive or non-immersive virtual reality was utilized. The number of citation left after this was six. For all six publications, an additional forward and backward search was carried out to ensure that no additional literature was missed.

The following literature review is based on a total of six publications (see Table 1). The relatively low number of selected publications in this specialized domain allowed us, in addition to just provide summaries and interpretation of study results, to make suggestions about what can be learned from the systems, the methodologies applied, and the results obtained.

5 LITERATURE REVIEW

5.1 Overview and Type Description
The six studies selected are all investigating the effect of VR training with regard to human factors considerations (see Table 1). Four of the studies include a search and rescue task in an urban environment (i.e., an indoor space), and two studies investigate aerial firefighting. Three of the studies are concerned with the training of direct firefighting tasks. The two studies by Clifford et al. (2018b,a) are dealing with the training of aerial attack supervisors who coordinate attack aircrafts for aerial firefighting, and the study by Cohen-Hatton and Honey (2015) deals with the training of commanders for urban scenarios.

5.2 Results

5.2.1 Search and Rescue
The studies by Bliss et al. (1997), Backlund et al. (2007), and Tate et al. (1997) were grouped together as they all investigate urban/indoor search and rescue scenarios. Bliss et al. (1997) focused on navigational training in a building with a preset route within a VR environment, using an HMD and a mouse for movement input, and contrasted this with either no training at all or with training the memorization of the route using a blueprint of the building. All three groups were subsequently assessed in a real building with the same layout as the training materials. The participants were told to execute a search and rescue in this building, with the two trained groups being advised to take the route that was trained prior. As expected, both the VR and blueprint training groups outperformed the group that received no prior training, as measured by completion time and navigation errors made. No
The results show a measurable improvement in the performance of firefighters using a VE for mission rehearsal over the other firefighters.

While Bliss et al. (1997) were more concerned with the fidelity of simulating a real scenario (without augmenting the content in any way), Backlund et al. (2007) specifically aimed to create a motivating and entertaining experience to increased training adherence, while eliciting physical and psychological stress factors related to a search and rescue task; they made use of game elements, such as score and feedback. Participants were divided into two groups, with one group receiving two training sessions using the VR simulation (called Sidh) before executing the training task in a real-world training area. The second group first performed the task in the training area and then did a single training session in the VR simulation. The VR environment was constructed by projecting the environment on four screens surrounding the participant. The direction of the participant was tracked, and movement enabled by accelerators attached to the boots (enabling walking in place as a locomotion input). The participants were tasked with carrying out a search and asked to evacuate any victims they came across. A score was displayed at the end of the attempt. Physical skills, such as body position and environment scanning, were tracked to allow for feedback mechanisms. The researchers found the simulation greatly increased learning outcomes, stating that performance in the simulation was better than traditional training method.
significantly better in the second session compared to the first. They highlight that the repeated feedback obtained during the first sessions resulted in a clear learning effect, which made participants more thorough in their second search a week later. Additionally, the tracking of the body position of participants, and relating appropriate feedback, resulted in the majority keeping a low position during the task, i.e., applying a vital safety skill. According to qualitative data, physical stress was elicited successfully. In addition, more than two thirds of the participants stated that they learned task relevant knowledge or skills. Participants generally stated that the simulation was fun.

The third study investigated the training of a search and rescue task in a novel environment, namely that of a Navy vessel (Tate et al., 1997). While not a traditional search and rescue task, i.e., the task was concerned with locating and extinguishing the fire while navigating the interior correctly, the general nature of the task, traversing an indoor environment for fire fighting supervisors (AASs). AAS units deployed in wildfire scenarios are tasked with coordinating attack aircraft that aim to extinguish and control the fire. These commanders are flying above the incident scene in a helicopter and need to assess the situation on the ground to coordinate fire attacks. The researchers put commanders in a simulated environment, showing a local wildfire scenario, using either a high-definition TV, an HMD (Oculus Rift CV1), or a CAVE setup (270° cylindrical projection).

5.2.2 Commander Training

Rather than assessing the execution of physical skills in VR, Cohen-Hatton and Honey (2015) evaluated the training of cognitive skills of commanders in a series of experiments. In their three-part study, the aim was to evaluate whether goal-oriented training, i.e., the evaluation of goals, the anticipation of consequences, and the analysis of potential risks and benefits for a planned action, would lead to better explicit formulation of plans and the development of anticipatory situational awareness. This was compared to groups given standard training procedures for the same scenarios. The researchers used three different scenarios as follows: a house fire, a traffic accident, and a fire threatening to spread across different buildings in an urban area. Participants encountered all three scenarios: first in a VR environment (experiment 1) and then on the fireground (experiment 2). Lastly, the house fire was recreated in a live-burn setting for the third experiment. Participants were compared based on whether they had received standard training or goal-oriented training procedures. The scenarios presented the participants with situations that demanded decisions to be taken dynamically based on new information that would be presented during the trial (e.g., an update of the location of a missing person, the arrival of a new fire crew, or sudden equipment failure). Their behavior was coded to obtain the frequency and chronology of occurrence of information gathering (situation assessment (SA)), plan development (plan formulation (PF)), executing a plan by communicating actions (plan execution (PE)), and anticipatory situational awareness. The researchers concluded that the VR environment accurately mirrors the commander activities as executed in real-life scenarios, because the chronology of SA, PF, and PE follows the same pattern for the group that received standard training. The patterns obtained during experiment two and three further support the notion of VR as a viable analog to real-life training. The behavior for the participants receiving goal-oriented training was further consistent across all degrees of realism, which supports the viability of VR for commander training.

The viability of training commanders utilizing immersive VR technology was also demonstrated by Clifford et al. in two studies (Clifford et al., 2018a; Clifford et al., 2018b). These studies complement the work carried out by Cohen-Hatton and Honey (2015), since the work environment and the nature of the measures were different while the overall question of the viability of a virtual environment for firefighter training remained the same. The first study (Clifford et al., 2018b) was investigating the effect of different types of immersion, by varying the display technology used, and their impact on the situational awareness of aerial attack supervisors (AASs). AAS units deployed in wildfire scenarios are tasked with coordinating attack aircraft that aim to extinguish and control the fire. These commanders are flying above the incident scene in a helicopter and need to assess the situation on the ground to coordinate fire attacks. The researchers put commanders in a simulated environment, showing a local wildfire scenario, using either a high-definition TV, an HMD (Oculus Rift CV1), or a CAVE setup (270° cylindrical projection).

While there were no differences in the abilities to accurately ascertain where the location of the fire is between display types, the location of secondary targets, such as people and buildings, was easier to determine with the HMD and CAVE setup which was attributed to the wider field of view (FOV) of these two display devices. The comprehension of the situation and the prediction of future outcomes, as part of the situational awareness scales, were also significantly better with the immersive VR options. The researchers found no significant differences between the two immersive display types for any of the subscales of the situational awareness measure. The researchers conclude that the immersive displays offer better spatial awareness for training firefighters in VR and are overall preferred by trainees compared to the non-immersive training.

The second study by Clifford et al. (2018a) investigated the elicitation of stress by manipulating interference in communication between the air attack supervisor and the pilots of the air attack aircraft. The AASs were put into a simulator that visualized a local wildfire using a CAVE setup (Figure 2). The AAS could communicate with the pilot of the helicopter sitting in and using the internal communication device hand signals, while using a foot pedal to activate...
outgoing radio communication with attack pilots and operations management. Communication disruptions were varied, first only using vibration of the seat (simulated in the CAVE) and the sound of the helicopter, then introducing background radio chatter from other pilots, and lastly, interrupting the radio transmissions to simulate a signal failure. Heart-rate variability and breathing rate were used as physiological measures of stress as well as self-report questionnaires for stress and presence were applied. The researchers conclude that the system was successful in simulating the exercise as all participants completed the task successfully. The trainees felt present in the virtual space, although the realism and involvement measured did not significantly differ from the observable midpoint. While the signal failure did not show a significant increase in physiological stress compared to the radio chatter condition, overall the physiological stress measures showed an increase in stress responses. It has to be noted that the researchers do associate the increase in breathing rate to the overall increase in communication between conditions and therefore discount this as a viable stress measure. Qualitative data, together with the self-report data, suggest that the communication disruption successfully induced stress in participants. The participants additionally reported enjoyment in using the system.

6 DISCUSSION

The studies reviewed for this article, despite limited numbers, do offer valuable insights into the viability of VR as a tool for firefighter training. Immersive VR technology provides an ecologically valid environment that mimics that of real-life exercises adequately. As shown by Clifford et al. (2018b), the use of monitor-bound simulations has limitations that negatively impact situational awareness. Being able to train spatial and situational awareness with a FOV that more closely resembles that of normal human vision, using an HMD or CAVE setup enables the creation of training environments in which trainees feel present. The studies conducted by Cohen-Hatton and Honey (2015) provide even stronger evidence for this, by showing that the behavior of their participants was consistent across levels of fidelity:

“In Experiments 1–3, the same scenarios were used across a range of simulated environments, with differing degrees of realism (VR, fireground, and live burns). The patterns of decision making were remarkably similar across the three environments, and participants who received standard training behaved in a manner that was very similar to that observed at live incidents […]”

While only applicable to two of the studies, the training of physical skills could successfully be done in the studies using natural input methods, by either tracking body posture or using firefighting gear as input devices. Trainees, when being provided with feedback in the virtual reality environment, do learn from their mistakes and improve the execution of physical skills in successive trials. This underscores the value of experiential learning enabled by VR. Natural input methods are becoming more and more prevalent for VR applications, due to the improvements in tracking. Two of the studies reviewed were conducted in the late 90s (Bliss et al., 1997; Tate et al., 1997), which resulted in constraints for the possibilities of more natural input. With both studies having been conducted more than 20 years ago as of the writing of this article, the outlook for future work by Bliss et al. (1997) was already anticipating the reappraisal of VR capabilities for training:

“The benefits of VR need to be assessed as the type of firefighting situation changes and as the capabilities and most of VR changes.”

On the other hand, the study conducted by Cohen-Hatton and Honey (2015) was concerned with commander training and therefore relied more heavily on decision-making tasks rather than physical skills, which are more easily simulated since the execution is mainly verbal.
Many of the studies observed, both old and new, make an effort to provide an ecologically valid environment for their simulation that is as analogous to the real-life activity as possible; even if, as previously stated, they are limited by technology. For example, both Backlund et al. (2007) and Cohen-Hatton and Honey (2015) required their participants to wear firefighting uniforms during their tasks (Figure 3). Bliss et al. (1997) did not require the participants to equip any firefighting gear but did give them industrial goggles sprayed with white paint to inhibit their vision in a similar manner how smoke would in a real scenario. Likewise, Backlund et al. (2007) use a real fire hose (Figure 4) to give a more apt input method than the joysticks and VR controllers used in the other studies observed.

However, there still remains much room for future research into furthering the ecological validity of the virtual environment within the context of firefighting training. Of all the studies observed, very few attempt to involve the senses outside of the auditory and visual systems. The inclusion of extra senses into the virtual environment—for example, haptic feedback (Hoffman, 1998; Insko et al., 2001; Hinckley et al., 1994) or smell (Tortell et al., 2007)—has been shown to improve learning outcomes, aid user adaptation to the virtual environment, and increase presence. Many studies already exist that can be incorporated into firefighting training to provide a richer and more realistic environment for the trainee. For example, Shaw et al. (2019) presented a system to replicate the sensation of heat (via heat panels) and the smell of smoke into their virtual environment [for smell, see also the FiVe FiRe system (Zybura and Eskeland, 1999)]; although in the context of a fire evacuation instead of firefighting training, the authors note that their participants demonstrated a more realistic reaction when presented with a fire. Likewise, Yu et al. (2016) present a purpose-built firefighting uniform that changes its interior temperature in reaction to the simulation. For haptic feedback, there is promising research into haptic fire extinguisher technology that could also be incorporated (Seo et al., 2019). Looking to commercial systems, the evaluation of other input methods could be promising for increasing ecological validity and improving the possible transfer of physical skills; see, for example, Flaim Trainer3 or Ludus VR4. While current studies (Jeon et al., 2019) are already showing promise in improving the ecological validity of firefighting training by partially incorporating these suggestions, additional research and study would be very beneficial to the field.

Regarding the training of mental skills, the review obtained ample evidence for the viability of skill transfer from VR to real deployment. Especially navigation tasks, requiring trainees to apply spatial thinking, were successfully trained in three of the reviewed articles. Training with VR was on par with memorizing the building layout utilizing blueprints and improves performance in subsequent real navigation tasks. As was highlighted by participants in the study by Tate et al. (1997), the VR training enabled spatial planning and subsequently improved performance:

“Most members of the VE training group used the VE to actively investigate the fire scene. They located landmarks, obstructions, and possible ingress and egress routes, and planned their firefighting strategies. Doing so enabled them to use their firefighting skills more effectively.”

Another important finding is the heightened engagement of trainees during VR training. The majority of studies reviewed found evidence for trainees preferring, enjoying, and being engaged with the training. The study by Backlund et al. (2007) went one step further by utilizing score and feedback systems to enhance engagement, which they deem to be important for voluntary off-hour usage of such a system. VR, as opposed to traditional training, provides the possibility of logging performance, analyzing behaviors, and providing real-time feedback to trainees without the involvement of trainers. Just as important as the frequency of training for the upkeep of skills, which is made possible with the relative ease of administration and the heightened engagement during VR training, the mental preparation of firefighters also plays an important role in counteracting possible adverse mental effects brought upon by threatening conditions during deployment. Physiological measures used by Clifford et al. (2018a) show that stress can be elicited successfully in a VR training scenario. Multi-sensory stimulation seems to further add to the realism and stress experienced as was stated in their study:

“With distorted communications and background radio chatter, you’re fearing to miss teammates communications and wing craft interventions. But the engine sound and the vibrations make the simulation much more immersive.”

Unlike many other studies in this inquiry, Bliss et al. (1997) concluded that the results from the group that used VR training

3https://www.flaimsystems.com/
4http://www.ludus-vr.com/en/portfolio/firefighter-services/
were not significantly better than their peers who used the traditional training solution—in this case, the use of blueprints. While the VR group performed on par with those who used blueprints, the results are underwhelming in comparison to other studies observed in this inquiry. In line with Engelbrecht et al. (2019), who deemed technology acceptance a weakness of VR technology in their analysis, the authors point to their participants’ low acceptance of technology and their familiarity with using the traditional training method as an explanation. However, while it is true that the study was conducted in times with arguably less prevalent acceptance of technology in general, this factor of familiarity, acceptance, and embrace of technology as a viable training tool should be considered in future work.

In addition to technological acceptance of potentially impacting learning outcomes, it is important to note the limitations of the technology used in all articles observed, especially earlier examples, and what effect this could have had on their results. Resolution of screens, their refresh rate, and the FOV of the headset have all improved significantly since the late 90s when two of the studies of this inquiry took place (see Table 2). Likewise, as Table 2 shows, earlier modern examples of HMDs, such as the Oculus Rift DK1, are considerably more under-powered than their more modern iterations.

As can be seen, the FOV of the headsets used in the older studies was significantly more constrained than any used in more recent research. The I-Glasses by virtual I/O used by Bliss et al. (1997) had only a 30-degree field of view in each eye while VR4 by virtual research systems used by Tate et al. (1997), a similarly aged study, had a FOV of 60°. For comparison, the more modern Oculus DK1 and CV1, used by Clifford et al. (2018a) and Cohen-Hatton and Honey (2015), have a FOV of 110°. This is potentially significant as, in the context of “visually scanning an urban environment for threats”, Ragan et al. (2015) found the participants performed substantially better with a higher FOV. Toet et al. (2007) found that limiting the FOV significantly hindered the ability of their participants to traverse a real-life obstacle course—a setting closer, albeit not virtual, to the task set by Bliss et al. (1997) and Tate et al. (1997). This limitation could potentially give further explanation as to why the VR training group did not outperform the blueprint group in the study of Bliss et al. (1997). However, in the study of Tate et al. (1997), the VR group outperformed traditional methods despite sharing the same limitation of FOV; although it is possible, as Ragan et al. (2015) suggest, that the limited FOV had other negative consequences, such as causing the user to adopt an unnatural method of moving their head to observe the virtual environment.

In addition, the lower refresh rates of some HMDs could be cause for consideration. Low refresh rates of headsets have been directly correlated to the sensation of cybersickness in VR (LaViola, 2000) which in turn has been shown in previous studies to significantly negatively affect the performance of participants in VR simulators (Kolasinski, 1995). For comparison, the Oculus CV1, as used by Clifford et al. (2018b), has a refresh rate of 90hz whereas the I-Glasses, VR4, and Oculus Rift DK1 as used by Bliss et al. (1997), Tate et al. (1997), and Cohen-Hatton and Honey (2015) can only produce a maximum of 60hz (Ave and Clara, 1994; Herveille, 2001) with Tate et al. (1997) specifying that their simulation ran at approximately 30 frames per second. As a baseline, LaViola (2000) note that:

“A refresh rate of 30 Hz is usually good enough to remove perceived flicker from the fovea. However, for the periphery, refresh rates must be higher.”

Therefore, all HMDs used in this inquiry, despite their age, should be within these limits. Bliss et al. (1997), with the lowest refresh rate of all studies observed, support this by stating that, unlike previous research, there was no sign of performance decrements in their study due to cybersickness with only two of their participants reported having experienced it. Likewise, Tate et al. (1997) used 1 minute rest breaks to avoid any simulation sickness which therefore mitigates any potential impact this would have had on their results. In addition, Cohen-Hatton and Honey (2015) report that only two of 46 of their participants experienced cybersickness despite the comparatively low refresh rate of the Oculus Rift DK1. However, it is important to note that various studies have shown that cybersickness affects female users more acutely than males (LaViola, 2000; Munafo et al., 2017), and in each of the aforementioned studies, the majority of the participants were male (Bliss et al. 1997 and Cohen-Hatton and Honey 2015): all participants were male; Tate et al. (1997): 8/12 participants were male. Therefore, any effect of negative impact on performance that could have been caused by

---

**Table 2 | A comparison of VR headsets.**

| Name            | Release date | Resolution (per eye) | Refresh rate                  | FOV (diagonal) |
|-----------------|--------------|----------------------|--------------------------------|----------------|
| VR4             | 1994         | 480 x 240            | NTSC timing (30/60 hz)        | 60             |
| I-Glasses       | 1995         | 300 x 200            | Maximum of 60 hz (VGA connection) | 30 per eye    |
| Oculus Rift DK1 | 2014         | 640 x 800            | 60 hz                         | 110            |
| Oculus Rift CV1 | 2016         | 1080 x 1200          | 90 hz                         | 110            |
| Valve Index     | 2018         | 1440 x 1660          | 90/120/144 hz                 | 120            |
| Vive Pro        | 2018         | 1440 x 1660          | 90 hz                         | 110            |
| HP Reverb G2    | 2020         | 2160 x 2160          | 90 hz                         | 1140           |
| Oculus Quest 2  | 2020         | 1832 x 1920          | 80 hz                         | 110            |
the lower refresh rates of the HMD may have been avoided—or, at least, mitigated—due to the gender distribution heavily leaning towards males in the firefighting profession (Hulett et al., 2007, 2008) which was reflected in the participant selection of the studies observed. Regardless, we can note that the refresh rates of all HMDs observed would not seem to detract from their findings, although future studies should attempt to use HMDs with a high refresh rate to avoid any such complications.

Both Tate et al. (1997) and Bliss et al. (1997) used a Silicon Graphics Onyx computer using the Reality Engine II to create the virtual environments. Likewise, Backlund et al. (2007) used the half-life 2 engine (released in 2004). While both engines were powerful for the time, computer hardware has increased exponentially since either of their releases (Danowitz et al., 2012). As such, these simulations have a much lower level of detail, both of the environment and the virtual avatar, than the more modern examples examined which use modern engines (such as Unity3D). This could potentially have an effect on the results from these studies and is important to investigate.

Regarding model fidelity’s effect on presence, Lugrin et al. (2015) found that no significant differences could be found between realistic and non-realistic environments or virtual avatars. Ragan et al. (2015), in the context of a visual scanning task, noted that visual complexity—which could include model/texture detail, fog, or number of objects in the environment—had a direct effect on task performance. Principally, they noted that the participants performed better in environments with fewer virtual objects. Due to this, Ragan et al. (2015) recommended that designers should attempt to match the visual complexity of the virtual environment to that of its real-life counterpart. However, the authors concede that different factors of visual complexity could affect the task performance in varying levels of severity and that future work would be required to gauge the impact of each factor. LUKosch et al. (2019) stated that the low physical fidelity of environments does not significantly impact learning outcomes or the ability to create an effective learning tool. Therefore, while there could be certain factors that are impacted by lower graphical quality, we cannot find sufficient grounds to discount or significantly question the results of the aforementioned studies.

7 CONCLUSION

While this review can only draw limited conclusions with regard to the viability of VR technology for general firefighter training, the scrutiny applied to the sourcing of publications provides an important step forward. The findings from previous work highlight the potential of VR technology to be an ecologically valid analog to real-life training in the acquisition of physical and mental skills. It can be applied to the training of commanders as well as to support the training of navigation tasks for unknown indoor spaces. The limitations of the technology used in the summarized studied, such as not being able to create and display high-fidelity immersive environments and the lack of using natural input methods, can be overcome with the developments that have been made in the immersive VR space over the past years. This opens up new opportunities for researchers to investigate the effectiveness of VR training for the target population. VR research for firefighters is wide open and promising, asEngelbrecht et al. (2019) stated in their SWOT analysis of the field: “Without adequate user studies, using natural input methods and VR simulations highly adapted to the field, there is little knowledge in the field concerning the actual effectiveness of VR training.”

While there is room to transfer findings from other domains to inform designs, evidence for the effectiveness of training itself should be approached with caution when drawing conclusions for the entirety of the emergency response domain. The work presented in this article can serve as a helpful baseline to inform subsequent research in this domain and might also be useful to inform the design of systems in adjacent domains; however, evidence of the effectiveness of training itself should not be generalized to other emergency response domains.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The research was conceptualized by HE and SH. HE drafted the first manuscript. SH supervised and guided the structure of the work, provided feedback, and revised iterations of the manuscript prior to submission.

FUNDING

This work was partially funded by Fire and Emergency New Zealand.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2021.671664/full#supplementary-material
