Journal of Safety Research 75 (2020) 205–221

Contents lists available at ScienceDirect

Journal of Safety Research

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

# Implementing and evaluating novel safety training methods for construction sector workers: Results of a randomized controlled trial

Mikko Nykänen *, Vuokko Puro, Maria Tiikkaja, Henriikka Kannisto, Eero Lantto, Frans Simpura, Jose Uusitalo, Kristian Lukander, Tuula Räsänen, Tarja Heikkilä, Anna-Maria Teperi

*Finnish Institute of Occupational Health, P.O. Box 40, FI-00032 Helsinki, Finland*

ARTICLE INFO

*Article history:*
Received 20 December 2019
Received in revised form 3 August 2020
Accepted 30 September 2020
Available online 21 October 2020

*Keywords:*
Virtual reality
Human factors safety training
Safety self-efficacy
Safety locus of control
Safety motivation

ABSTRACT

*Introduction:* The construction industry is regarded as one of the most unsafe occupational fields worldwide. Despite general agreement that safety training is an important factor in preventing accidents in the construction sector, more studies are needed to identify effective training methods. To address the current research gap, this study evaluated the impact of novel, participatory safety training methods on construction workers' safety competencies. Specifically, we assessed the efficacy of an immersive virtual reality (VR)-based safety training program and a participatory human factors safety training program (HFST) in construction industry workplaces. *Method:* In 2019, 119 construction sector workers from eight workplaces participated in a randomized controlled trial conducted in Finland. All the study participants were assessed using questionnaires at baseline, immediately after the intervention and at one-month follow-up. We applied generalized linear mixed modeling for statistical analysis. *Results:* Compared to lecture-based safety training, VR-based safety training showed a stronger impact on safety motivation, self-efficacy and safety-related outcome expectancies. In addition, the construction sector workers who participated in the VR-based safety training showed a greater increase in self-reported safety performance at one-month follow-up. Contrary to our study hypotheses, we found no significant differences between the study outcomes in terms of study participants in the HFST training condition and the comparison condition without HFST training. *Conclusion:* Our study indicates that VR technology as a safety training tool has potential to increase safety competencies and foster motivational change in terms of the safety performance of construction sector workers. In the future, the efficacy of participatory human factors safety training should be studied further using a version that targets both managerial and employee levels and is implemented in a longer format. *Practical implications:* Safety training in virtual reality provides a promising alternative to passive learning methods. Its motivating effect complements other safety training activities.



## 1. Introduction

The construction industry has a high risk of occupational accidents and injuries and is regarded as one of the most unsafe industries worldwide (Pinto, Nunes, & Ribeiro, 2011; Zhou, Goh, & Li, 2015). Previous studies (Pinto, Nunes, & Ribeiro, 2011; Guo et al., 2012) highlight the lack of occupational safety training as one of the major factors behind the poor level of occupational safety in the construction industry. The importance of safety training for enhancing workers' safety competencies has also been emphasized in previous review articles (Ricci et al., 2016; Robson et al., 2012).

Glendon et al. (2006) highlight several elements of occupational safety training that can significantly improve its effectiveness. These include the use of photographs and demonstrations, as well as the active participation of workers and various simulations of work situations. Participatory and engaging learning methods in particular have been found to be more effective than more passive safety training approaches (e.g., lectures; Burke et al., 2006, 2011). Participatory safety training is based on active participation on the part of learners and utilizes reflection and dialogue as a means of knowledge and skills development (Burke et al., 2006). Despite general agreement that safety training is an important factor for preventing accidents in the construction sector, we need more studies to identify effective training methods (see van der Molen et al., 2018).

* Corresponding author at: Finnish Institute of Occupational Health, PO Box 40, FI-00032, Finland.

*E-mail address:* mikko.nykanen@ttl.fi (M. Nykänen).

https://doi.org/10.1016/j.jsr.2020.09.015
0022-4375/

To address the current research gaps, the present randomized-controlled trial (RCT) assessed the impact of novel and participatory safety training methods on construction workers' safety competencies. Specifically, we assessed: (a) the efficacy of an immersive virtual reality-based safety training program, and (b) the efficacy of a participatory human factors safety training program in construction industry workplaces.

The purpose was to provide information that can be used in the construction industry when considering choices in terms of safety training methods. This RCT was pre-registered and the rationale, hypotheses and methods of this study were defined previously (Anonymous, 20xx). This study focused on the efficacy and process evaluation of intervention processes, and its report followed CONSORT guidelines (Schulz et al., 2010).

### 1.1. VR-based safety training

In the general sense, virtual reality (VR) refers to a computer-generated artificial 3D environment (Li et al., 2018). New technologies such as virtual reality offer tools for developing engaging methods of safety training. Specifically, VR provides potential for emphasizing the active role of the learner during safety training activities (Lee, Wong, & Fung, 2010). A basic distinction has been made between immersive VR and non-immersive VR (see Plechatá et al., 2019). Immersion refers to the extent to which a virtual environment is capable of delivering an inclusive, extensive, and vivid illusion to the human senses (Slater & Wilbur, 1997). A key difference between non-immersive and immersive VR is related to the level of interaction between the user and the VR environment. Non-immersive VR is often based on technologies such as computer desktops or tablet screens. In these cases, the virtual environment is viewed through a computer screen and people use a keyboard to control their movements in the virtual environment. Thus, users can see the contents of the virtual environment but have limited capabilities to interact with the environment. Immersive VR utilizes more advanced technologies such as VR headsets and systems that track head and body movements. In immersive VR, the user interacts with the computer-generated environment by physically moving and manipulating objects using motion-tracking devices. Although VR as a concept is not particularly new in safety training, the use of immersive VR technology in safety training is not yet widespread (Grabowski & Jankowski, 2015). The development of VR hardware and software is making immersive VR-based safety training a feasible option for delivering workplace safety training. Safety learning based on immersive VR has received a considerable amount of attention in the construction sector in recent years. (Li et al., 2018).

Empirical evidence to support the effectiveness of VR-based safety training is rather limited (see Gao et al., 2019). Also, evidence concerning the benefits of VR-based safety training is limited and has often focused on students, apprentices, or young workers. The few previous studies available have shown mixed results. Sacks, Perlman, and Barak (2013) found that immersive VR safety training had a beneficial impact on safety-related knowledge acquisition among construction workers with limited work experience and civil engineering students. In their RCT study, the participants used a VR headset and a hand-held controller. Using a non-randomized study design, Gheisari and Esmaeli (2019) found that VR-based safety training using a computer desktop platform enhanced the hazard recognition skills of construction sector trainees. In a randomized-controlled trial by Leder et al. (2019), participants used a VR headset and a wireless controller to navigate and interact in the VR environment. Using an apprentice student sample, they found that safety training outcomes in risk perception did not differ according to whether the training had been immersive VR-based safety training or training using PowerPoint slides. In our study, we focus on applications of immersive virtual reality. Therefore, in the following chapters, VR safety training refers to training based on immersive VR technology, which allows users to interact with the computer-generated environment.

### 1.2. Human factor perspective in safety training

Another relatively new concept in safety science is related to applying the holistic human factors perspective in safety training. This perspective is important to acknowledge in safety training processes because human behavior plays a role in most workplace accidents (Ford & Tetrick, 2008). Furthermore, Teperi, Puro, and Ratilainen (2017) and Teperi et al. (2015) have shown that learning to recognize human factors at work has positive outcomes for safety skills. However, few studies have explored the outcomes of safety training based on human factors in construction sector workplaces. In our study, we implemented a human factors safety training program with a participatory and holistic approach (HFST training). The holistic approach means that the safety training addresses the role of human activity at various levels: the individual level (e.g., attitudes), group level (e.g., information flow), and the organizational level (e.g., resources). HFST training guides workers towards identifying the promoting or hindering effect of human factors in safety.

## 2. Conceptual background

Our evaluation framework utilized social cognitive theory (Bandura, 1977, 1997; Rotter, 1966, 1982) and the theory of planned behavior (Ajzen, 1991). These psychological theories have been used previously in safety research by Guerin et al. (2019), Casey, Krauss, and Turner (2018), and Nykänen, Sund, and Vuori (2018). The key concepts of our evaluation framework included safety self-efficacy, safety locus of control, outcome expectancies, safety knowledge, and safety motivation. Previous empirical studies have shown that these individual-level factors play an important role in safety performance and safety outcomes (e.g. accidents) (Christian et al., 2009; Jones & Wuebker, 1993; Katz-Navon et al., 2007; Neal & Griffin, 2006; Arezes & Miguel, 2008). Thus, they are important targets of safety training efforts among the construction sector workforce. Next, we describe the key concepts of our evaluation framework.

### 2.1. Individual-level safety competencies

Knowledge of preventive action is a strong determinant of safety behavior at the workplace (Griffin & Neal, 2000). According to a definition by Christian et al. (2009), safety knowledge refers to knowing how to perform work safely (e.g., safety procedures, using protective equipment). However, increasing safety knowledge is not enough to impact on behavior. Individuals must acknowledge their personal role in safety, identify the benefits of a particular preventive action, and be convinced that they can successfully perform the required safety behaviors (see DeJoy, 1996; Jones & Wuebker, 1993). Safety self-efficacy refers to the degree of confidence in one's ability to perform safety-related activities successfully, such as identifying hazards, recognizing factors that affect the occurrence of accidents, and using personal protective equipment (see Bandura, 1997; Nykänen, Sund, & Vuori, 2018). According to previous studies (Bandura, 1997), verbal persuasion (e.g., positive feedback) and mastery experience, in which an individual successfully practices a behavior, are key mechanisms for enhancing self-efficacy. Internal safety locus of control refers to the perception that one's own work behavior plays an important role in preventing injuries or harmful events. In contrast, external safety locus of control is related to the belief that accidents happen due to external factors such as luck or the actions of other people. (Jones & Wuebker, 1993). A similar concept, "perceived control over safety issues" refers to the degree to which people perceive themselves as affecting safety improvements at the workplace (Curcuruto, Mearns, & Mariani, 2016). Safety self-efficacy, safety locus of control, and perceived control over safety issues are closely interrelated psychological constructs. Nevertheless, there are a few distinctions. The safety locus of control measure is related to personal beliefs concerning accidents, and the perceived control over safety issues is more related to day-to-day safety-related activities at work (Curcuruto, Mearns, & Mariani, 2016). Furthermore, safety self-efficacy refers to the degree of confidence in one's ability to perform a specific safety-related task or activity and safety locus of control is an attribution toward safety outcomes in general. It refers to perceptions of the location in which one's control over accidents or injuries resides. For example, an individual may acknowledge that accidents are dependent on their behavior but may have low self-efficacy for specific preventive actions (Nykänen et al., 2019).

By the concept of safety-related outcome expectancies, we refer to the expected value that workers give to safety-related activities such as the use of personal protective equipment. Arezes and Miguel (2008) used a similar concept of outcome value in studying predictors of the use of hearing protection devices. Previous studies have also emphasized the role of safety motivation as a proximal antecedent for safety performance at workplaces (Christian et al., 2013). According to Neal, Griffin, and Hart (2000), safety motivation refers to "an individual's willingness to exert effort to enact safety behaviors and the valence associated with those behaviors". Furthermore, safety performance refers to actions or behaviors that individuals perform to promote health and safety (Burke, Scheuer, & Meredith, 2007). Previous studies have divided it into subdimensions that include safety compliance and safety participation. Safety compliance refers to core safety activities that need to be carried out by individuals (e.g., using personal protective equipment) and safety participation refers to participating in voluntary safety activities (Neal, Griffin, & Hart, 2000).

### 2.2. Participants' responsiveness to safety training

Previous research has highlighted the importance of examining implementation processes as part of safety intervention evaluation (van der Molen, 2007; see Dugan & Punnet, 2017). Exploring intervention delivery adds explanatory value to the efficacy evaluation (see Carroll et al., 2007). Participant responsiveness has been defined as an important criterion for evaluating implementation processes. According to Berkel et al. (2011), participant responsiveness refers to engagement, active participation, interest, and satisfaction related to the intervention program. In this context, we incorporated an investigation of various aspects of participant responsiveness in our process evaluation.

The VR-based learning process has its own special features, which we took into account when evaluating the implementation process. Immersion and "sense of presence" have been identified as key features of VR-based training (Li et al., 2018). Sense of presence has been defined as the psychological consequence of immersion. It refers to the degree of feeling connected and involved, absorbed, and engaged in a given virtual environment (McCreery et al., 2013; Slater & Wilbur, 1997; Lombard & Ditton, 1997). Thus, the sense of presence reflects responsiveness to, psychological engagement in and satisfaction with VR-based safety learning activities. Moreover, Mantovani and Castelnuovo (2003) highlighted that the sense of presence is a key feature ensuring the efficacy of VR-based training. We evaluated four components of the sense of presence as active ingredients of a VR learning experience: (a) Spatial presence, referring to the individual's perception of being located in an environment that is conveyed by VR technology; (b) Involvement, defined as the psychological state of focusing one's energy and attention on activities and events in the virtual environment; (c) Realism, referring to the perceived consistency of the information conveyed by a VR with real-world experience; and (d) Feeling of control, that is, the perception of control a person experiences over the task environment or when interacting with the virtual environment (Witmer & Singer, 1998; Hartmann et al., 2015).

One deficit of VR learning activity is its potential risk of producing simulator sickness symptoms. Simulator sickness is a syndrome similar to motion sickness (Dużmańska, Strojny, & Strojny, 2018). According to previous studies, simulator sickness symptoms are common during VR experiences. (see Salzman et al., 1999), although it is rare that these sickness symptoms are severe or strong (Bouchard, Robillard, & Renau, 2007; Lawson et al., 2002). Nevertheless, simulator sickness can limit the efficacy of VR-based training and should be monitored (see Dużmańska, Strojny, & Strojny, 2018).

## 3. Study aims and hypotheses

The first aim of this study was to investigate the efficacy of VR-based safety training in comparison to that of trainer-centered lecture-based safety training. The second aim was to investigate the impact of the participatory HFST training. Third, we applied process evaluation to determine whether the intervention activities were implemented as intended. The purpose was to provide insights into the delivery process of the training processes and to assess target group responsiveness. The VR-based safety training and HFST training were both carried out as part of the same RCT conducted in Finnish construction companies.

In accordance with previously published study methods (Nykänen et al., 2020), our hypotheses were the following: In comparison to trainer-centered and lecture-based safety training, we expected VR-based safety training to have a stronger effect on safety competencies, defined as: (1) safety self-efficacies, (2) safety locus of control, (3) safety knowledge, (4) safety-related outcome expectancies, (5) safety motivation, and (6) safety performance. In addition, we hypothesized that the participatory HFST training would have a positive impact on (7) safety self-efficacies and (8) safety-related outcome expectancies and (9) safety knowledge.

## 4. Study methods

### 4.1. Study participants

This RCT was implemented in eight Finnish construction sector workplaces in 2018–2019. To ensure a sufficient number of participants per organization, only medium- and large-sized organizations were recruited. Participation in the study required the target organizations to make arrangements regarding employees' work schedules and organize their working time to enable them to participate in the intervention processes. The research team first identified potential Finnish construction sector companies that were considered to have the required personnel resources to participate in the research project. The key inclusion criteria here concerned the size of the organization and having several years of experience in the construction sector field. All the listed companies also had previous experience in occupational safety development projects. These factors were expected to support the implementation process and the success of the study's data collection. During the recruitment of the workplaces, safety managers were sent an email containing an invitation to participate in a research project and informational material regarding the study. If the safety managers were interested, a meeting with the workplace representatives was arranged to present the study process and intervention methods. In all, 21 construction sector workplaces were approached during the recruitment process. A meeting was arranged with a total of eight organizations after initial contact (email and phone call). Eight construction sector organizations were interested in the research project and agreed to participate. When the organization decided to participate in the study, an employee meeting was held to disseminate information on the study and to recruit participants. Initially, the inclusion criteria for participation in the study were the following: (a) age between 20 and 50, (b) at least two years of work experience in the construction sector, (c) Finnish as a native language, and (d) current employment as a construction worker. Eligible employees who expressed interest in the study were invited to an information presentation by a research team member. In the employee meeting, the research team members described the study and collected both written consent forms for participation and the completed baseline questionnaires. All the participants were volunteers.

Due to recruitment difficulties encountered with the first two target organizations, we re-examined our inclusion criteria and changed the age criteria so that those over 50 years of age were also admitted to the study. Thus, in terms of age, the participant inclusion criteria were changed to at least 20 years of age. Altogether, 120 eligible construction sector workers gave their informed consent for the study and participated in the baseline measurements.

Our study had three measurement time points (baseline, short-term follow-up, and one-month follow-up). All the study measurements were taken at the workplaces of the participants, where the research team member handed the paper format questionnaire to the participants. In one of the target organizations, the baseline questionnaires were completed electronically during a Skype employee meeting. The number of workers participating in the study varied among the organizations from 9 to 20 ($M = 15$, $SD = 3.25$). Both VR safety training and lecture-based safety training took place at the workplaces approximately two weeks after baseline measurements. The HFST training was also implemented at the workplaces. Here, the implementation process included one exception: In one target organization, employees worked in different cities. For practical reasons, all the safety training sessions for this organization were held at one location.

### 4.2. Randomization process and follow-ups

A statistician, who was not involved in contacting the target organizations or the data collection procedures, used a computer-generated randomization method to classify the study participants into one of four intervention arms after baseline measurements. Stratified randomization was used to establish equal intervention conditions in terms of work experience. Randomization was stratified by work experience: 2–5 years, 6–15 years and over 15 years. It was conducted separately for each participating organization. The study participants who were allocated into Intervention arm 1 participated only in the VR-based safety training. Intervention arm 2 participated only in lecture-based safety training. Intervention arm 3 participated in both the VR-based safety training and HFST training, and Intervention arm 4 participated in both lecture-based safety training and HFST training. Given the nature of the safety training programs, it was not possible to blind study participants to their condition allocation.

In total, 60 study participants were randomized to participate in the VR-based safety training and 60 to participate in the lecture-based safety training. The participants were informed whether they were allocated to the VR-based training or lecture-based training approximately one week before the training. Furthermore, 59 study participants were randomized to also participate in the HFST training one month after the VR-based or lecture-based training. The study participants allocated to the HFST training were informed approximately two weeks before the training. Sixty-one participants were assigned to the control condition without HFST training.

At short-term follow-up immediately after the VR-based training and lecture-based training (T2), 110 participants (92.4%) completed the questionnaire. Respectively, 105 participants (88.2%) completed the questionnaire at one-month follow-up (T3). The study participants allocated to the HFST training completed the questionnaire immediately after their training. The control group without HFST training completed the questionnaire at the equivalent times.

We found no statistically significant differences in study outcomes between the respondents who provided follow-up data (at T2 and T3) and those who provided only baseline measures. Thus, we considered that the data were missing at random. No-show bias was controlled for by collecting follow-up data on all the study participants originally randomized and included in the study. Of the 60 participants assigned to the VR-based safety training condition, 3.3% did not participate in the training activities but provided follow-up data at T2, and 10% of the 60 participants randomized to lecture-based safety training did not participate in the lecture-based safety training but provided data at T2. Furthermore, of the 59 participants assigned to participate in the HFST training, 6.7% did not participate in HFST training activities but provided follow-up data at T3. In accordance with the intention-to-treat principle, we based our analyses on complete randomized intervention data, which included both those who participated in the training activities and nonparticipants. During data screening, we removed one multivariate outlier from the analysis due to extreme negative scores in the study variables and potential misreporting. Detailed information on the study flow and recruitment process is outlined in Fig. 1.

### 4.3. Interventions

A detailed description of the intervention processes has been provided in previous studies (Nykänen et al., 2020). Due to space limitations, we provide only a brief overview of the VR-based safety training, lecture-based safety training, and HFST training evaluated in this study.

#### 4.3.1. VR-based safety training

The VR environment was developed using the Unity3D game engine, C# programming language, and HTC Vive VR system. The content of the VR-based safety training included: (1) Acknowledging the positive outcomes of using personal protective equipment; (2) Identifying and preventing hazards in a work environment (falling objects, slipping and tripping); (3) Hazard awareness around traffic on a construction site; (4) Inspecting equipment and machinery for potential defects or hazards; and (5) Safety procedures for using a table saw or safety procedures for working at height. To ensure that the content of the training was compatible with the needs of the participating workplaces, we developed two different versions of the training. The only difference between the two versions was in their final part (5). The first version (implemented in four study organizations) focused on the safe use of a table saw, and the other on safe working procedures at height (implemented in four study organizations). Nevertheless, both versions had equivalent training topics in terms of the final part of the training: inspecting equipment and identifying hazards in the work environment. Appendix A shows illustrative picture of the VR learning environment. Internal pilot testing of the VR-based

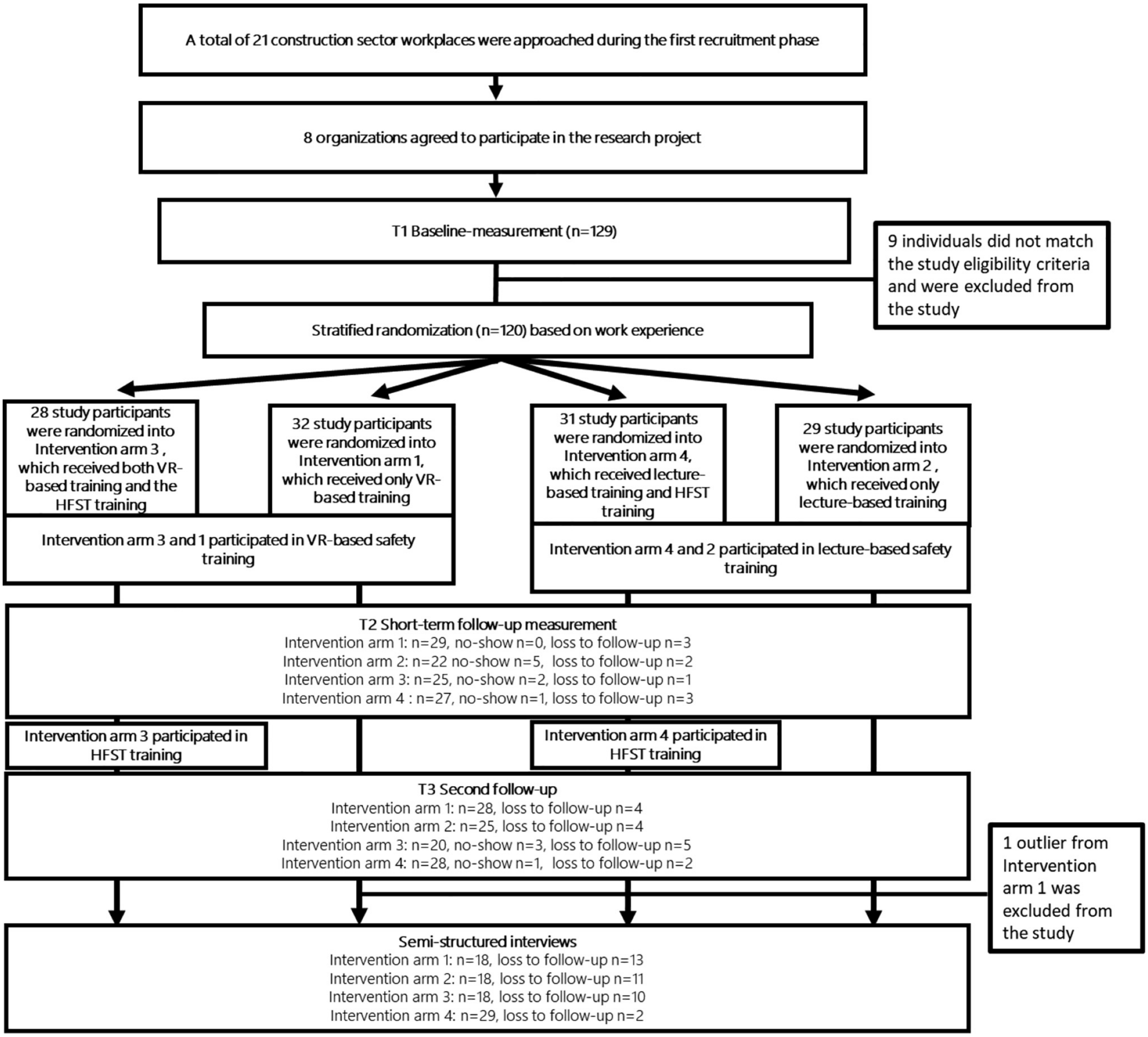



**Fig. 1.** Description of enrollment, allocation and measurement points in the study.

training was conducted from December 2018 to January 2019. A total of 14 people from outside the research team tested the VR learning environment. This was to ensure technical and pedagogical functionality of the training before the actual implementation.

The VR learning activities used social cognitive theories (Bandura, 1997) and the theory of planned behavior (Ajzen, 1991). For example, safety self-efficacy was targeted by providing mastery experiences of preventive actions and sharing positive feedback (visual and auditory) on the safety-related actions performed by the study participants. Internal safety locus of control was targeted by guiding study participants to identify the factors preceding accidents and raising awareness of how an individual worker's preventive actions can prevent accidents. For example, the study participants practiced preventative actions by marking the factors contributing to tripping or slipping accidents. Outcome expectancies were modified by providing study participants with feedback on the relationship between safety-related actions and accident prevention. The duration of the VR-based safety training was approximately 25–30 min. The virtual environment was developed so that it provided guidance for the study participants, which meant that the research team members did not need to provide any additional guidance.

### 4.3.2. Lecture-based safety training

The lecture-based safety training was based on a trainer-centered approach. The learning topics and the content of the lecture-based training corresponded to the VR-based training sessions implemented in the study organization, and the distinguishing factor was the training method (VR vs. lecture). A safety training expert who was not part of the research team delivered the lecture-based safety training at each workplace separately. The trainer was actively involved in lecturing, while the study participants were in a more passive learner role. The active participation of the study participants (e.g., engaging in discussions) during the lectures was neither specifically prevented nor promoted. The safety-related information was presented in a slide-show format in which screenshots from the VR safety training were shown. In this respect, our intervention process was similar to that in previous study by Leder et al. (2019). Each lecture was attended by 3–9 study participants ($M$ = 5.4 participants). Social cognitive theories

(Bandura, 1997) and the theory of planned behavior (Ajzen, 1991) were considered when planning the lectures. A key mechanism for enhancing self-efficacy was verbal persuasion and encouraging the participants to implement safety-related actions at their workplaces. Safety locus of control was targeted by sharing information about the importance of individual workers' preventive actions. Outcome expectancies were modified by providing the study participants with information on the relationship between safety-related actions and accident prevention. The duration of the lecture-based safety training was also approximately 25–30 min.

*4.3.3. HFST training*

The HFST training was tailored to meet the needs of the construction sector. Two professional safety trainers planned the training session together and both used the predefined protocol in their training. The training process supported active participation on the part of study participants (e.g., reflection and dialogue, short participatory group/pair discussions). The trainer guided the study participants to identify the positive outcomes of acknowledging human factors in promoting safety and to identify how human factors can contribute to safety. At the end of the training, the participants analyzed a fictive accident case and identified the factors that either weakened or strengthened safety during the accident. This participatory activity was a key mechanism for providing the trainees with mastery experiences. HFST training can also be seen as complementary to lecture / VR training. HFST training provided an overview of the human factors and a broader perspective of the factors affecting safety and accidents. Case examples were used, which were linked to the content of VR/lecture training. The total duration of the HFST training was approximately 90 min, including a 10 min break. Appendix B contains a description of the educational topics of the HFST training.

*4.4. Outcome measures*

The outcome measures were included in the questionnaires at all measurement points (baseline, short-term follow-up, one-month follow-up). An exception was the self-reported safety performance measures, which were included at both baseline and one-month follow-up.

*4.4.1. Safety motivation*

Safety motivation was measured using three 7-point items (1 = Strongly disagree, 7 = Strongly agree) adapted from Neal, Griffin, and Hart (2000). Cronbach's $\alpha$ was 0.66 at T1, $\alpha = 0.68$ at T2, and $\alpha = 0.65$ at T3. An example item is: "*I feel that it is worthwhile putting an effort into maintaining or improving my personal safety*."

*4.4.2. Safety locus of control*

Four 7-point items (1 = strongly disagree, 7 = Strongly agree) adapted from the study by Mazaheri Hidarnia and Ghofranipour (2012) were used to measure safety locus of control. The scale assessed beliefs regarding accident causation and had two subdimensions, each with two items. The first dimension measured the internal safety locus of control (e.g., *people can avoid injury by being careful and aware of potential dangers*) and the second dimension measured the external safety locus of control (e.g., *Whether people get injured or not is a matter of fate, chance or luck*). The Spearman–Brown reliability estimate was 0.54 (item correlation 0.37, $p < 0.00$) for internal safety locus of control at T1, 0.44 (item correlation 0.28, $p < 0.00$) at T2, and 0.73 (item correlation 0.56, $p < 0.00$) at T3. The Spearman–Brown reliability estimate was 0.41 (item correlation 0.28, $p < 0.00$) for external safety locus of control at T1, 0.62 (item correlation 0.45, $p < 0.00$) at T2, and 0.73 (item correlation 0.41, $p < 0.00$) at T3. This indicated that the safety locus of control measure had low reliability. Therefore, the respective analyses should be regarded as only indicative.

*4.4.3. Perceived control over safety issues*

We measured the perceived control over safety issues using two 7-point items (1 = strongly disagree, 7 = Strongly agree) (e.g. *workers are able to make significant contributions to the safety of the work area*) adapted from the study by Curcuruto et al. (2016). The reliability estimate was 0.76 (item correlation 0.61, $p < 0.01$) at T1, 0.86 (item correlation 0.76, $p < 0.01$) at T2, and 0.71 (item correlation 0.56, $p < 0.01$) at T3.

*4.4.4. Safety knowledge*

Our safety knowledge measure was a modified version of the safety knowledge test used previously by Sacks, Perlman, and Barak (2013). Our purpose was to measure whether study participants correctly identified hazards and safety procedures during a work task (using a table saw or working at height). The study participants completed the safety knowledge measurement at baseline, immediately after the VR- or lecture-based training, and at one-month follow-up. The safety knowledge measure included a brief description of a work task with three open-ended questions: (1) List the potential hazards related to a described work task; (2) Describe how to ensure occupational safety at the beginning of the work task; (3) List the safety procedures, appropriate safety equipment and safety instructions and actions that help prevent accidents during the work task.

The data preparation for the safety knowledge measure included the following steps: The research team prepared a comprehensive list of hazards and safety procedures related to the open-ended questions. Second, the content validity of the list was confirmed by two independent safety training experts working in the construction sector. Third, the final list was used as frame to score the safety knowledge level of the study participants. One research team member, blinded to the intervention arms, scored open-ended questions by comparing the study participants' responses to the comprehensive list of hazards and safety procedures. The participants received one point for each relevant hazard, safety procedure, piece of safety equipment, and preventive action mentioned. Finally, a safety knowledge sum score (0–24 points) was calculated for each study participant. The research team member who scored the answers was not involved in the intervention development, did not participate in the baseline or follow-up assessments, and had no contact with the study participants or workplaces.

*4.4.5. Safety self-efficacies and outcome expectancies*

Two measures, safety self-efficacies and safety-related outcome expectancies were developed for this study. To establish the content validity of these measures, we used the following steps: (1) The questionnaire items were developed by the research team; (2) The construction sector workers ($n = 3$) reviewed the measures and checked that the questionnaire items were meaningful to the target group. These construction sector workers did not participate in the study and did not work in the target organizations.

The 14-item safety-related outcome expectancies scale assessed the perceived benefits (1 = not at all useful 7 = extremely useful) in terms of individual-level and organizational-level safety-related activities. The 10-item safety self-efficacy scale assessed the study participants' confidence (1 = very poorly 7 = very well) in various safety-related activities, including work task-specific safety performance and more general safety behaviors (e.g., identifying hazards).

According to Bandura (2012), self-efficacy is a multidimensional construct and self-efficacy scales should be linked to relevant activity domains. Previous studies have also utilized multidimensional outcome expectancy measures (e.g. Wójcicki, White, and McAuley, 2009). Thus, we treated both safety-related outcome expectancies and safety self-efficacy measures as multidimensional constructs.

The safety self-efficacy measures included the following subdimensions: (1) three items assessed self-efficacy to identify the factors affecting safety (example item: "*Identify the issues affecting the occurrence of accidents,*" $\alpha = 0.86$ at T1, $\alpha = 0.80$ at T2, $\alpha = 0.85$ at T3); (2) four items assessed self-efficacy related to preventive actions (example item: *"Use personal protective equipment,"* $\alpha = 0.73$ at T1, $\alpha = 0.64$ at T2, $\alpha = 0.77$ at T3); and (3) three items assessed work task-specific safety self-efficacy (example item: "*Carry out safety inspections of table saw,*" $\alpha = 0.85$ at T1, $\alpha$ 0.89 at T2, $\alpha$ 0.90 at T3).

The safety-related outcome expectancy measure included the following sub-dimensions: (1) six items measured the worker-level performance outcome value (example item: "*Worker compliance with safety instructions,*" $\alpha = 0.89$ at T1, $\alpha = 0.89$ at T2, $\alpha = 0.90$ at T3); (2) four items measured the organizational level practices outcome value (example item: *"Increasing the amount of safety training at the workplace,"* $\alpha = 0.82$ at T1, $\alpha = 0.80$ at T2, $\alpha = 0.82$ at T3); and (3) four items assessed the general safety performance outcome value (example item *"Ensuring cleanliness and tidiness at the workplace,"* $\alpha = 0.82$ at T1, $\alpha = 0.80$ at T2, $\alpha = 0.82$ at T3). A detailed description of the questionnaire items and the psychometric properties of the outcome expectancy and safety measures are presented in the Supplementary material.

*4.4.6. Safety performance*

We assessed safety performance using a self-report approach. We adapted a scale previously used by Neal, Griffin, and Hart (2000): A four-item seven-point scale (1 = strongly disagree, 7 = Strongly agree) which included two subdimensions: (1) safety compliance ("*I use the correct safety procedures to do my job,*" "*I use all the necessary safety equipment to do my job*"); and (2) safety participation ("I *voluntarily perform tasks or activities that help improve workplace safety,*" "*I put in extra effort to improve the safety of the workplace*"). The Spearman–Brown reliability estimate was 0.79 (item correlation 0.65, $p < 0.00$) for safety compliance at T1, and 0.71 (item correlation 0.55, $p < 0.00$) at T3. The reliability estimate was 0.70 (item correlation 0.54, $p < 0.00$) for safety participation at T1, and 0.73 (item correlation 0.57, $p < 0.00$) at T3.

*4.5. Intervention measures*

The intervention condition variable was coded with a value of 1 for study participants who took part in the VR-based safety training and 0 for study participants who took part in lecture-based safety training. In the analyses, which evaluated the effects of HFST training, the intervention condition variable was coded with a value of 1 for study participants who took part in the HFST training and 0 for study participants who had no HFST training.

*4.6. Implementation process measures*

We utilized observational data, data recorded during activities in the VR environment, and study participants' self-reports of learning experiences to evaluate various aspects of the intervention processes.

*4.6.1. Recorded VR data*

We recorded the research participants' movements in a VR learning environment to gain insights into the level of activity of the study participants during the training. Specifically, we used the 3D position and rotation data of the VR headset. Data processing used Python programming. In VR, the user views the virtual world through two separate cameras (one for each eye) so the best approximation for the position of the user was the midpoint of these two cameras, which we used to gather the position data. Thus when we refer to user movement, we specifically mean the movement of the user's head. Using this location information, we produced an illustration of how users navigate through the VR environment during a safety training session.

*4.6.2. Sense of presence*

We measured the sense of presence as an indicator of participant responsiveness to VR-based safety training. We used a seven-point measure adapted from previous studies by Schubert, Friedmann, and Regenbrecht (2001) and Witmer, Jerome, and Singer (2005). Our measure included four sub-dimensions, with higher scores indicating greater sense of presence: (1) Spatial presence included three items (e.g. "*I felt present in the virtual space,*" 1 = strongly agree 7 = strongly disagree) with $\alpha = 0.91$; (2) The involvement sub-dimension was reduced from three to two items because one reverse-coded item ("*How aware were you of the real world surrounding you*?") did not correlate with the other items. Thus, the involvement sub-dimension used in the study contained two items (e.g. "*How much did the virtual environment involve you*?" 1 = not at all 7 = very much) with a Spearman–Brown coefficient of 0.63 and item correlations of 0.46 ($p < 0.01$). 3) Judgements of realness included three items (e.g. "*How real did the virtual world seem to you*?" 1 = not at all 7 = very much) with $\alpha = 0.82$ and 4) Feelings of control included four items (e.g. "*How much were you able to control events in the virtual environment*?" 1 = not at all 7 = very well) with $\alpha = 0.64$.

*4.6.3. Simulator sickness symptoms*

In our study, we used a 16-item measure by Kennedy et al. (1993) to measure simulator sickness symptoms immediately after the VR-based training. Each item reflected simulator sickness symptoms (e.g., blurred vision, dizziness, increased salivation) and was rated by the participants as "0 = none," "1 = slight," "2 = moderate," or "3 = severe." Thus, the simulator sickness measure score could range from 0 to 48. In our study, Cronbach's $\alpha$ for simulator sickness measures was 0.58.

*4.6.4. Perceived utility and satisfaction*

Satisfaction with the training and perceived utility were assessed at short-term follow-up (after VR-based safety training and lecture-based safety training) using thirteen 7-point items. Six items measured whether the study participants perceived that the training had enhanced their safety-related skills and abilities (1 = not at all, 7 = very much). Seven items measured the participants' general satisfaction with the training. Satisfaction with the training and perceived utility was also assessed at one-month follow up among those who participated in the HFST training. Two 7-point questionnaire items assessed the perceived benefits of HFST training, and two 7-point items measured general satisfaction with the training. A detailed description of the perceived utility and satisfaction items is provided in Table 5 and in the Results section.

*4.6.5. Observational data*

To further document the intervention delivery, we used a structured checklist during the VR-based safety training, lecture-based safety training, and HFST training. Observational data were collected by a research team member and were based on direct observation of the training sessions. In the VR-based safety training, observational coding was targeted toward the following factors: (1) The amount of study participant guidance during the VR activities, (2) adjustments of the VR headset during training, and (3) Technical issues during training. In the lecture-based safety training and HFST training, observation gathered information on (1) whether participants contributed their own information to the discussions, (2) the level of discussion between the study participants, and (3) the number of comments the participants made during the training. The observational data from each session was entered into a spreadsheet and scored quantitatively from 1 (not at all) to 5 (very much).

#### 4.6.6. Interview data

After the one-month follow-up, short telephone interviews were conducted to gather feedback on the training sessions. The study participants were given the opportunity to share their experiences of the training in their own words. The interviews lasted approximately five minutes each and were voice recorded. The interview data were processed as follows. First, the data were divided into different categories based on which safety training method they referred to (VR, lecture or HFST). Next, distinct anecdotes or stated opinions about learning experiences were identified and categorized. The emerging categories were discussed, agreed on, and determined by five research team members. These categories were further classified as positive or negative feedback. Once the final code list was agreed upon, one of the coders coded the transcripts. At this point, the frequencies of the negative and positive feedback were calculated to obtain an idea of their prevalence in the data. Thus we calculated the percentage of the participants who provided positive and negative feedback on each safety training method.

### 4.7. Statistical methods

#### 4.7.1. Descriptive statistics and process analysis

The frequencies, correlations, and means of the study variables were calculated to provide information on the study participants, descriptive statistics, and implementation process-related factors. Baseline differences were tested using independent samples t-tests and chi-square tests according to the scale of the variable. Outlier detection was performed using visual inspection of data and by identifying three or more standard deviations from the mean in all the outcome variables (Osborne & Overbay, 2004). Cronbach's alpha coefficient was calculated to determine the internal consistency of the scales with three or more items. The Spearman–Brown reliability estimate and the Pearson correlation coefficient were calculated using two-item scales (see Eisenga, Grotenhuis, & Pelzer, 2013).

#### 4.7.2. Outcome analyses

As the study participants were nested in different organizations, the individuals' responses were potentially influenced by their workplaces. Furthermore, outcome measures were collected over time, producing a design in which measurement sessions were nested within individuals. In our study, we used generalized linear mixed modeling (GLMM) to evaluate safety training outcomes. Time was modelled as a repeated variable and an unstructured covariance structure was used to accommodate dependencies between different points of time. All analyses were adjusted for age and organization-level random intercepts were included in all the models to account for nesting.

Our primary study hypotheses focused on the differential changes in the VR-based safety training participants and the lecture-based safety training participants over short-term follow-up and one-month follow-up. Time was treated as a variable with two levels (baseline and follow-up) and the intervention condition as having two levels (VR-based safety training, lecture-based safety training). Our secondary research questions focused on the differential changes in the HFST training condition participants and the study participants without HFST training between the short-term follow-up and the one-month follow-up. In these analyses, the intervention condition was treated as having two levels (HFST training, no HFST training). For all outcome analyses, the intervention effect of interest was the interaction between intervention condition and time.

Significance levels were set at $p < 0.05$ and Satterthwaite small-sample correction was used. The normality of the residuals in all the study models were confirmed before analyses. We calculated Cohen's d effect sizes using the method described by Morris (2008). Effect sizes for between-group differences were interpreted based on Cohen's (1988) guidelines, with 0.20 indicating a small effect, 0.50 indicating a medium effect, and 0.80 indicating a large effect. Following the intention-to-treat principle, the study participants were analyzed according to the intervention condition to which they were randomly assigned, regardless of whether they attended the training.

## 5. Results

### 5.1. Descriptive statistics

Table 1 describes the study participant characteristics across the intervention conditions. The mean age of the study participants was 41.2 years, and the sample was 96.6% male. The study participants randomly allocated into the VR-based safety training or lecture-based safety training did not differ in terms of their background variables. In addition, we detected no statistically significant baseline differences between study participants randomized into the HFST training and those randomized into the comparison group without HFST training.

Furthermore, we found no statistically significant baseline differences between intervention conditions (VR-based vs. lecture-based) in terms of study outcomes and we detected no statistically significant differences at short-term follow-up between the study participants further allocated to HFST training and those without HFST training. Tables 2 and 3 provide an overview of the mean scores of the study participants' outcome variables at baseline, short-term follow-up, and one-month follow-up. The intercorrelations of the study variables are provided in the Supplementary material.

### 5.2. Short-term effect of virtual reality-based safety training

Next, we present the analyses in terms of the short-term effects of the VR-based safety training. The first set of GLMM models included only the time variable (assessment time points: baseline, short-term follow-up) as a fixed effect. These models were designed to assess the impact of time in general, for the entire study sample. We found that all the study outcomes, except external safety locus of control and perceived control over safety issues, increased significantly from baseline to short-term follow-up, as indicated by the models in which only time was a predictor.

Next, to explore the different time trends at short-term follow-up between the intervention conditions, we added an interaction term between the intervention condition and the time variable in the models. These models included the time variable (baseline, short-term follow-up), the intervention condition (0 = lecture-based safety training, 1 = VR-based safety training), the interaction of the intervention condition and time, and age as fixed effects. We hypothesized that VR-based safety training would have a stronger impact on study outcomes at short-term follow-up. This assumption was partly supported. Our analyses showed a statistically significant interaction between time and the intervention condition in five outcomes: Compared to the lecture-based safety training, the study participants in the VR-based safety training showed a greater increase in self-efficacy in identifying factors affecting safety (esti-

**Table 1**
Study participant characteristics.

| Variable | Total sample $n = 119$ Mean (SD) [min–max] | Study participants in VR-based training $n = 59$ Mean (SD) [min–max] | Study participants in lecture-based training $n = 60$ Mean (SD) [min–max] | Study participants in HFST training $n = 60$ Mean (SD) [min–max] | Study participants without HFST training $n = 59$ Mean (SD) [min–max] |
|---|---|---|---|---|---|
| Age | 41.2 (10.9) [21–64] | 41.8 (11.7) [21–64] | 40.7 (10.1) [24–61] | 41.9 (10.9) [21–64] | 40.5 (10.9) [23–64] |
| Work experience in current occupational field | 15.0 years (11.9) [2–41] | 15.2 years (11.2) [2–41] | 13.8 years (9.3) [2–37] | 15.2 years (10.3) [2–37] | 13.6 years (10.2) [2–41] |
| Has had work-related accident | | | | | |
| Yes | 30.8% | 26.8% | 35.5% | 28.8% | 31.7% |
| No | 69.2% | 73.3% | 65.0% | 71.2% | 68.3% |
| Job tenure | 10.4 years (9.10) [0.5 months–37 years] | 10.6 years (10.1) [1 month–37 years] | 10.1 years (8.0) [0.5 months–37 years] | 11.8 years (9.6) [0.5 months–37 years] | 8.9 years (8.3) [2 months–37 years] |
| Personal previous usage of VR devices | | | | | |
| – Never used | 67.8% | 75.9% | 60.0% | 66.1% | 69.5% |
| – Tried | 27.1% | 19.0% | 35.0% | 28.8% | 25.4% |
| – Use once a month | 1.7% | 3.4% | 0% | 0% | 3.4% |
| – Use once a week or more often | 3.4% | 1.7% | 5.0% | 5.1% | 1.7% |
| Working as a supervisor | | | | | |
| – Yes | 10.2% | 13.8% | 6.7% | 11.9% | 8.5% |
| – No | 89.8% | 86.2% | 93.3% | 88.1% | 91.5% |
| Previous safety training *(1 = not at all 7 = very much)* | 3.8 (1.6) [1–7] | 4.0 (1.6) [1–7] | 3.6 (1.6) [1–7] | 3.89 (1.6) [1–7] | 3.78 (1.6) [1–7] |

**Table 2**
Mean scores and standard deviations of outcome measures by intervention condition (VR-based vs. lecture-based).

| Outcome | Baseline Mean (sd) [min–max] | | Short-term follow-up Mean (sd) [min–max] | | Long-term follow-up Mean (sd) [min–max] | |
|---|---|---|---|---|---|---|
| | VR-based safety training ($n = 59$) | Lecture-based safety training ($n = 60$) | VR-based safety training ($n = 56$) | Lecture-based safety training ($n = 55$) | VR-based safety training ($n = 51$) | Lecture-based safety training ($n = 54$) |
| General safety self-efficacy | 5.93 (0.63) [4.25–7.00] | 5.88 (0.59) [4.25–7.00] | 6.06 (0.57) [5.00–7.00] | 6.03 (0.49) [4.75–7.00] | 6.16 (0.59) [5.00–7.00] | 6.03 (0.56) [5.00–7.00] |
| Self-efficacy in identifying factors affecting occupational safety | 5.75 (0.67) [3.67–7.00] | 5.92 (0.69) [4.00–7.00] | 6.09 (0.54) [4.67–7.00] | 5.94 (0.58) [4.67–7.00] | 6.10 (0.58) [4.67–7.00] | 5.84 (0.61) [4.33–7.00] |
| Work task-specific safety self-efficacy | 5.74 (1.02) [2.67–7.00] | 5.41 (1.17) [1.67–7.00] | 5.98 (0.82) [3.00–7.00] | 5.78 (0.83) [3.00–7.00] | 5.99 (0.84) [3.67–7.00] | 5.74 (0.92) [2.67–7.00] |
| Worker-level performance outcome value | 6.51 (0.55) [4.67–7.00] | 6.58 (0.54) [5.00–7.00] | 6.73 (0.43) [5.33–7.00] | 6.67 (0.50) [4.83–7.00] | 6.57 (0.53) [4.50–7.00] | 6.59 (0.55) [4.33–7.00] |
| Organizational level practices outcome value | 5.90 (1.03) [3.25–7.00] | 5.73 (0.96) [2.75–7.00] | 6.10 (0.78) [4.00–7.00] | 5.76 (0.87) [3.50–7.00] | 6.10 (0.75) [4.25–7.00] | 5.81 (0.99) [3.00–7.00] |
| General safety performance outcome value | 6.64 (0.48) [4.50–7.00] | 6.56 (0.54) [5.00–7.00] | 6.75 (0.38) [5.00–7.00] | 6.70 (0.49) [4.75–7.00] | 6.67 (0.40) [5.50–7.00] | 6.68 (0.42) [5.50–7.00] |
| Safety motivation | 6.55 (0.62) [4.67–7.00] | 6.63 (0.59) [4.67–7.00] | 6.74 (0.49) [5.00–7.00] | 6.70 (0.59) [4.67–7.00] | 6.75 (0.48) [5.33–7.00] | 6.80 (0.31) [6.00–7.00] |
| Internal safety locus of control | 6.40 (0.69) [4.00–7.00] | 6.47 (0.60) [5.00–7.00] | 6.69 (0.54) [4.00–7.00] | 6.52 (0.58) [5.00–7.00] | 6.41 (0.65) [4.50–7.00] | 6.52 (0.62) [4.50–7.00] |
| External safety locus of control | 2.12 (1.15) [1.00–6.50] | 2.15 (1.06) [1.00–5.00] | 1.94 (1.20) [1.00–6.00] | 2.01 (1.11) [1.00–6.00] | 2.05 (1.23) [1.00–6.50] | 2.11 (1.01) [1.00–5.00] |
| Perceived control over safety issues | 6.44 (1.14) [1.00–7.00] | 6.49 (0.60) [4.00–7.00] | 6.58 (0.90) [1.00–7.00] | 6.46 (0.59) [5.00–7.00] | 6.57 (0.51) [5.00–7.00] | 6.39 (0.57) [5.00–7.00] |
| Safety knowledge | 6.18 (4.61) [0.00–21.00] | 6.35 (4.04) [0.00–21.00] | 8.35 (4.76) [0.00–24.00] | 7.12 (3.47) [0.00–15.00] | 6.57 (4.06) [0.00–17.00] | 7.59 (4.11) [0.00–20.00] |
| Safety compliance | 6.24 (0.74) [3.00–7.00] | 6.25 (0.74) [3.00–7.00] | – | – | 6.59 (0.44) [5.50–7.00] | 6.46 (0.52) [5.50–7.00] |
| Safety participation | 5.22 (1.24) [2.50–7.00] | 5.31 (1.20) [1.50–7.00] | – | – | 5.84 (0.91) [4.00–7.00] | 5.42 (1.12) [1.50–7.00] |

mate 0.25, $p < 0.05$), internal safety locus of control (estimate 0.25, $p < 0.00$), safety motivation (estimate 0.09, $p < 0.05$), worker-level performance outcome value (estimate 0.10, $p < 0.05$), and safety knowledge (estimate 1.35, $p < 0.05$). Cohen's d effect sizes were 0.47 for self-efficacy in identifying factors affecting safety, 0.37 for internal safety locus of control, 0.20 for safety motivation, 0.32 for safety knowledge, and 0.25 for worker-level performance outcome value. Thus, the relative advantage of the VR-based safety training compared to the lecture-based safety training was small to moderate at short-term follow-up. Table 4 presents the results of the models exploring the safety training method by time interactions at short-term follow-up.

**Table 3**
Mean scores and standard deviations of outcome measures by intervention condition (HFST training vs. without HFST training).

| Outcome | Short-term follow-up Mean (sd) [min–max] | | One-month follow-up Mean (sd) [min–max] | |
|---|---|---|---|---|
| | HFST training condition ($n = 55$) | No HFST training ($n = 56$) | HFST training condition ($n = 52$) | No HFST training ($n = 53$) |
| General safety self-efficacy | 6.02 (0.57) [5.00–7.00] | 6.07 (0.49) [4.75–7.00] | 6.09 (0.60) [5.00–7.00] | 6.09 (0.55) [5.00–7.00] |
| Self-efficacy in identifying factors affecting occupational safety | 5.99 (0.61) [4.67–7.00] | 6.04 (0.52) [4.67–7.00] | 5.88 (0.62) [4.33–7.00] | 6.05 (0.58) [5.00–7.00] |
| Work task-specific safety self-efficacy | 5.93 (0.71) [3.33–7.00] | 5.83 (0.94) [3.00–7.00] | 5.76 (0.85) [2.67–7.00] | 5.96 (0.92) [3.00–7.00] |
| Worker-level performance outcome value | 6.63 (0.50) [5.33–7.00] | 6.76 (0.42) [4.83–7.00] | 6.57 (0.52) [5.17–7.00] | 6.59 (0.56) [4.33–7.00] |
| Organizational level practices outcome value | 5.86 (0.85) [4.00–7.00] | 6.00 (0.83) [3.50–7.00] | 5.90 (0.78) [3.75–7.00] | 6.00 (1.00) [3.00–7.00] |
| General safety performance outcome value | 6.72 (0.43) [5.00–7.00] | 6.74 (0.45) [4.75–7.00] | 6.65 (0.43) [5.50–7.00] | 6.70 (0.38) [5.50–7.00] |

**Table 4**
Results of GLMM models. Models exploring short-term effect of time and intervention condition effect in terms of time on the outcome variables.

| Outcome variable | Models exploring effect of Time[a] | | | Models exploring intervention condition (VR vs. lecture-based safety training) by time interaction[b] | | |
|---|---|---|---|---|---|---|
| | Time fixed effect Estimate (Std.error) | CI 95 % | *p.* | Intervention condition × time interaction fixed effect Estimate (Std.error) | CI 95 % | *p.* |
| General safety self-efficacy | 0.14 (0.05) | **<0.01**, 0.241 | **<0.01** | −0.04 (0.05) | −0.144, 0.056 | **<0.05** |
| Self-efficacy to identify factors affecting occupational safety | 0.17 (0.05) | 0.076, 0.279 | **<0.01** | 0.25 (0.12) | 0.002, 0.515 | **0.04** |
| Work task specific safety-self-efficacy | 0.29 (0.04) | 0.212, 0.379 | **<0.01** | −0.14 (0.14) | −0.457, 0.161 | **<0.05** |
| Worker level performance outcome value | 0.15 (0.02) | 0.104, 0.212 | **<0.01** | 0.10 (0.05) | 0.010, 0.208 | **0.03** |
| Organizational level practices outcome value | 0.11 (0.03) | 0.046, 0.180 | **<0.01** | 0.12 (0.14) | −0.180, 0.431 | 0.40 |
| General safety performance outcome value | 0.12 (0.03) | 0.068, 0.187 | **<0.01** | −0.04 (0.05) | −0.146, 0.061 | <0.05 |
| Safety motivation | 0.13 (0.05) | 0.020, 0.246 | **<0.05** | 0.09 (0.04) | 0.007, 0.177 | **0.03** |
| Internal safety locus of control | 0.16 (0.04) | 0.081, 0.247 | **<0.01** | 0.21 (0.06) | 0.076, 0.346 | **<0.01** |
| External safety locus of control | −0.17 (0.10) | −0.387, 0.036 | 0.10 | −0.02 (0.26) | −0.553, 0.505 | 0.92 |
| Perceived control over safety issues | 0.06 (0.05) | −0.036, 0.162 | 0.20 | 0.16 (0.15) | −0.137, 0.461 | <0.05 |
| Safety knowledge | 1.47 (0.44) | 0.494, 2.449 | **<0.01** | 1.36 (0.59) | 0.166, 2.569 | 0.**02** |

[a] Fixed effects in model: Time (baseline, short-term follow-up) and age including random intercept for study organization.
[b] Fixed effects in model: Time (baseline, short-term follow-up), intervention condition (1 = VR-based safety training 0 = lecture-based safety training), Time × intervention interaction and age including random intercept for study organization.

### *5.3. Long-term effect of virtual reality-based safety training*

Next, we explored the effect of time on outcome variables at one-month follow-up. These models included only the time variable (baseline, one-month follow-up) as a fixed effect. The purpose was to assess the long-term impact of participating in the study in general, for the entire study sample. We found that six study outcomes increased significantly from baseline to one-month follow-up in the study sample.

In the following analyses, we evaluated the differential changes between the study participants in the VR-based safety training and those in the lecture-based safety training over the one-month follow-up. These models included the time variable (baseline, one-month follow-up), the intervention condition (1 = VR-based training 0 = Lecture-based training), the interaction of the intervention condition and time, and age as fixed effects. We found a statistically significant interaction between the intervention condition and time in two outcomes: Compared to the lecture-based safety training, the VR-based safety training participants showed a greater increase in self-efficacy in identifying factors affecting safety (estimate 0.38, $p < 0.00$). The VR-based safety training participants also showed a greater increase in self-reported safety performance (estimate 0.46, $p < 0.00$). This means that they reported a higher rate of performing voluntary activities that help improve workplace safety and put extra effort into improving workplace safety. Cohen's d effect sizes at long-term follow-up were 0.61 for self-efficacy in identifying factors affecting safety, and 0.42 for safety participation, which indicated a small to moderate relative advantage for VR-based safety training. Table 5 presents the results of the models exploring the interaction between the safety training method and time at one-month follow-up.

### *5.4. HFST training effects*

Approximately half of the study participants were further randomized to participate in HFST training after the short-term follow-up (T2). Our final set of outcome analyses focused on comparing changes in terms of self-efficacies and outcome expectancies from short-term follow-up to one-month follow-up in two conditions: the HFST training condition and the control condition without HFST training. We evaluated the HFST training effect using models that included the time variable (short-term follow-up and long-term follow-up), the HFST training variable (0 = no HFST training 1 = HFST training), and age as fixed effects. Table 6 presents the results of statistical models exploring the effects of HFST training. Contrary to our study hypotheses, there were no significant differences between the increase of safety self-efficacies or outcome expectancies of the study participants in the HFST training condition and those in the control condition without HFST training. This indicated that participating in HFST training had no impact on the measured outcomes.

**Table 5**
Results of GLMM models. Models exploring long-term effect of time and intervention condition effect in terms of time on outcome variables.

| Outcome variable | Models exploring effect of Time[a] | | | Models exploring intervention condition (VR vs. lecture-based safety training) by time interaction[b] | | |
|---|---|---|---|---|---|---|
| | Time fixed effect Estimate (Std.error) | CI 95 % | p. | Intervention × time interaction fixed effect Estimate (Std.error) | CI 95 % | p. |
| General safety self-efficacy | 0.15 (0.05) | 0.050, 0.265 | **<0.01** | 0.04 (0.06) | −0.084, 0.180 | 0.47 |
| Self-efficacy to identify factors affecting occupational safety | 0.12 (0.06) | −0.010, 0.250 | <0.01 | 0.38 (0.13) | 0.113, 0.647 | **<0.01** |
| Work task specific safety-self-efficacy | 0.25 (0.04) | 0.172, 0.329 | **0.00** | −0.01 (0.16) | −0.246, 0.210 | 0.87 |
| Worker level performance outcome value | 0.02 (0.03) | −0.038, 0.096 | 0.40 | 0.02 (0.09) | −0.174, 0.222 | 0.81 |
| Organizational level practices outcome value | 0.11 (0.05) | −0.003, 0.227 | <0.05 | 0.05 (0.18) | −0.316, 0.422 | 0.77 |
| General safety performance outcome value | 0.07 (0.03) | 0.005, 0.139 | **<0.01** | −0.10 (0.08) | −0.281, 0.068 | 0.22 |
| Safety motivation | 0.18 (0.05) | 0.071, 0.294 | **0.00** | 0.01 (0.07) | −0.127, 0.155 | 0.84 |
| Internal safety locus of control | 0.01 (0.04) | −0.071, 0.097 | 0.76 | −0.08 (0.10) | −0.297, 0.134 | 0.45 |
| External safety locus of control | −0.05 (0.12) | −0.305, 0.191 | 0.64 | −0.02 (0.26) | −0.534, 0.493 | 0.93 |
| Perceived control over safety issues | 0.01 (0.07) | −0.126, 0.153 | 0.85 | 0.23 (0.21) | −0.182, 0.654 | 0.26 |
| Safety knowledge | 0.79 (0.46) | −0.149, 1.733 | <0.01 | −0.80 (0.56) | −1.934, 0.328 | 0.16 |
| Safety compliance | 0.24 (0.06) | 0.146, 0.347 | **<0.01** | 0.12 (0.08) | −0.035, 0.287 | <0.01 |
| Safety participation | 0.31 (0.09) | 0.129, 0.510 | **0.00** | 0.46 (0.15) | 0.165, 0.759 | **0.00** |

[a] Fixed effects in model: Time (baseline, one-month follow-up) and age including random intercept for study organization.
[b] Fixed effects in model: Time (baseline, one-month follow-up), intervention condition (1 = VR-based safety training 0 = lecture-based safety training), Time × intervention interaction and age including random intercept for study organization.

**Table 6**
Results of GLMM models. Models exploring HFST training effect in terms of time on outcome variables.

| Outcome variable | Models exploring effect of Time[a] | | | Models exploring intervention condition (HFST training vs. no HFST training) by time interaction[b] | | |
|---|---|---|---|---|---|---|
| | Time fixed effect Estimate (Std.error) | CI 95 % | p. | HFST training X time interaction fixed effect Estimate (Std.error) | CI 95 % | p |
| General safety self-efficacy | 0.03 (0.03) | −0.040, 0.106 | 0.37 | 0.04 (0.07) | −0.093, 0.184 | 0.51 |
| Self-efficacy to identify factors affecting occupational safety | −0.05 (0.06) | −0.194, 0.075 | 0.37 | −0.15 (0.09) | −0.362, 0.043 | 0.11 |
| Work task specific safety-self-efficacy | −0.02 (0.03) | −0.091, 0.043 | <0.01 | −0.28 (0.16) | −0.630, 0.066 | 0.10 |
| Worker level performance outcome value | −0.11 (0.03) | −0.193, −0.036 | 0.00 | 0.09 (0.05) | −0.021 0.209 | 0.10 |
| Organizational level practices outcome value | 0.00 (0.02) | −0.066, 0.049 | <0.01 | −0.03 (0.11) | −0.261, 0.201 | 0.79 |
| General safety performance outcome value | −0.04 (0.01) | −0.080, −0.015 | 0.00 | −0.03 (0.02) | −0.084, 0.021 | 0.23 |
| Safety knowledge | −0.70 (0.41) | −1.525, 0.124 | 0.09 | 0.12 (0.76) | −1.407, 1.657 | 0.87 |

[a] Fixed effects in model: Time (short-term follow-up, one-month follow-up) and age including random intercept for study organization.
[b] Fixed effects in model: Time (short-term follow-up, one-month follow-up), intervention condition (1 = HFST training 0 = no HFST training), Time × intervention interaction and age including random intercept for study organization.

### *5.5. Process evaluation*

#### *5.5.1. Process analysis of VR-based and lecture-based safety training*

The aim of the process evaluation was to evaluate whether intervention processes were delivered as intended and to assess study participant responsiveness. First, we explored whether the learning experiences of the lecture-based and VR-based safety training differed. The results of the independent samples t-tests showed significant differences between VR-based safety training and lecture-based safety training. The VR-based safety training participants reported a statistically significantly ($p < 0.01$) higher perceived utility and satisfaction levels than the lecture-based safety training participants. The results of the perceived utility and satisfaction questions are presented in Table 7. The highest mean difference between the perceived utility of the lecture-based and VR-based safety training was associated with identifying factors that may cause danger at the workplace (4.79 vs. 3.41) and with abilities to make safety-related observations (5.00 vs. 3.69). This supported our efficacy results, suggesting that the VR-based training had a greater impact on self-efficacy in identifying factors affecting safety. The highest mean difference in terms of satisfaction was detected in relation to how inspiring the safety training was perceived as being (5.94 vs. 3.91). This supported our finding that the VR-based training had a stronger impact on safety motivation.

We examined the sense of presence scores to determine the level of participant responsiveness in terms of psychological engagement. The study participants rated their learning experiences according to four sub-dimensions of sense of presence: (1) spatial presence (mean = 6.03 SD = 0.9); (2) controllability (mean = 5.85 SD = 0.62); (3) involvement (mean = 5.47 SD = 1.04); and (4) realism (mean = 5.31 SD = 0.96). Considering the response scale was from 1 to 7, these results indicated relatively high levels of sense of presence.

The reported simulator sickness mean score was 2.77 (range from 0.00 to 32.0, SD = 3.49), which is a similar result to previous studies that have evaluated simulator sickness symptoms related to VR experiences (e.g., Merta & Kelling, 2015). Thus, the results did not indicate that the VR experience had strong adverse effects on the study participants. Recorded data of the research participants' movements in a VR learning environment illustrated that study participants were exploring the VR environment as part of their learning. Case example of study participants movements during training is presented in Appendix A.

Of the employees who participated in the VR-based training and responded to telephone interviews ($n = 36$), 81% gave positive

**Table 7**
Perceived utility and satisfaction related to VR-based safety training and lecture-based safety training.

| Process evaluation questionnaire items | VR-based safety training Mean (SD) [min–max] | Lecture-based safety training Mean (SD) [min–max] | p | 95% CI of mean difference |
|---|---|---|---|---|
| **Perceived utility** | | | | |
| ***Safety training enhanced your skills in... (1 = not at all 7 = Quite a bit)*** | | | | |
| Identifying hazards at the workplace | 4.79 (1.80) [1–7] | 3.41 (1.58) [1–6] | <0.01 | 0.71, 2.05 |
| Making observations concerning safety at work | 5.00 (1.67) [1–7] | 3.69 (1.53) [1–6] | <0.01 | 0.72, 2.04 |
| Carrying out preventive actions | 4.92 (1.74) [1–7] | 3.63 (1.61) [1–6] | <0.01 | 0.63, 1.95 |
| Affecting occupational safety when using a table saw[a] | 5.15 (1.92) [1–7] | 3.86 (1.62) [1–7] | <0.01 | 0.59, 1.99 |
| Making inspections to promote occupational safety when using a table saw[b] | 5.19 (1.87) [1–7] | 4.00 (1.45) [1–7] | <0.01 | 0.52, 1.85 |
| Ensuring safety at work | 5.06 (1.90) [1–7] | 3.86 (1.70) [1–7] | <0.01 | 0.48, 1.91 |
| **General satisfaction** | | | | |
| ***To what extent do you agree or disagree... (1 = Strongly disagree 7 = Strongly agree)*** | | | | |
| I found the training useful for improving occupational safety | 6.17 (0.97) [3–7] | 4.94 (1.63) [1–7] | <0.01 | 0.70, 1.75 |
| I believe that I will change one or more of my work habits due to the training | 4.57 (1.83) [1–7] | 3.65 (1.56) [1–6] | <0.01 | 0.24, 1.58 |
| The training was inspiring | 5.94 (1.06) [3–7] | 3.91 (1.61) [1–7] | <0.01 | 1.49, 2.56 |
| I learned new things during the training | 4.08 (2.06) [1–7] | 2.92 (1.65) [1–7] | <0.01 | 0.41, 1.89 |
| The training methods supported my learning | 5.47 (1.28) [1–7] | 4.09 (1.82) [1–7] | <0.01 | 0.76, 2.00 |
| I was focused during the training | 6.38 (0.63) [5–7] | 5.51 (1.58) [1–7] | <0.01 | 0.40, 1.34 |
| The training method hindered my ability to focus during training | 1.64 (1.41) [1–7] | 2.44 (1.42) [1–6] | <0.01 | −1.35, −0.23 |

[a] For the 55 study participants, the question format was: "*affecting occupational safety when lifting objects*".
[b] For the 55 study participants, the question format was: "m*aking inspections to promote occupational safety when lifting objects*".

feedback on the training. In this feedback, training was seen as having increased in-depth thinking about occupational safety. The participatory approach was also considered good and the training was perceived as interesting. The interviews revealed that after the training, overall cleanliness had improved and work machinery inspections had increased in many study organizations. The employees claimed it was good that during the training they had had to think about and carry out the safety-related tasks themselves. Nineteen percent of those interviewed also gave negative feedback. The negative issues mentioned were mostly related to the technical problems in implementing the training.

The results (presented in Table 8) of the observational data showed that 40.4% of the study participants had minor technical problems during the VR-based safety training. The majority (41%) of the observed technical issues were related to minor positional tracking issues with the motion controllers when they froze or began to drift away from their actual position in the real world. These were nevertheless fixed quickly. Furthermore, 14% of the observed technical issues were related to minor difficulties with performing the hand waving motion in the virtual environment in a way that the computer would recognize. Seven study participants stepped out of the play area during their VR-based training session. In these cases, a research team member redirected the person back to the play area. Other technical issues were related to minor software- or audio-related issues in the initial phase of the training. In all cases, the duration of the technical issues was short. An exception was the first two VR-based training sessions in the study, which had to be canceled in the first minutes. The research team scheduled new VR-based training times for these two study participants. Observational data also showed that the VR learning environment worked quite independently and that the research team did not need to give the research participants much additional guidance. Results of VR-based safety training observations are presented in Table 8.

Using *t*-tests, we compared the perceived utility, general satisfaction, and sense of presence scores of the study participants with technical difficulties ($n = 28$) during the VR-based safety training and those of the study participants without technical difficulties ($n = 26$). We found no statistically significant differences in any of the sense of presence subscales. However, our analysis revealed that the study participants with technical difficulties had higher ratings (2.07 vs. 1.11, $p < 0.05$) on items reflecting dissatisfaction with the training method ("*the training method hindered my ability to focus*"), indicating that the technical issues during the VR-based safety training had an impact on the learning experience. Overall, the process analysis results showed that the program was delivered as intended in terms of minimal risk of simulator sickness and immersive VR learning experience.

Our observational data also confirmed that the lecture-based safety training used a trainer-centered approach. The level of discussion between the study participants was low in the lecture-based safety training sessions (88.9% no discussion during training, 11.1% some discussion). The study participants themselves did not share much safety knowledge during the lectures (77.8% none, 22.2% some safety knowledge sharing). The overall amount of comments made by the study participants was also low during the lectures (55.6% none, 44.4% some comments during training). Of the employees who participated in the lectures and responded to the telephone interviews ($n = 47$), 51% gave positive feedback on the training. According to the feedback, the training provided an opportunity to rehearse safety-related topics. Furthermore, 56% also expressed negative feedback, according to which the lecture contained no new information. Sitting while training was also mentioned as a negative factor. The usual work tasks of construc-

**Table 8**
Observations during VR-based training ($n = 54$).

| Observation category | None | Minor | Average | Substantial | Extreme |
|---|---|---|---|---|---|
| Study participant was instructed in how to operate in the VR environment (after tutorial phase) | 30.2% | 34.0% | 22.6% | 3.8% | 9.4% |
| Study participant was instructed in how to use motion controllers during the training (after tutorial phase) | 83.0% | 9.4% | 7.5% | 0% | 0% |
| The VR headset was adjusted during the training (after the tutorial phase) | 98.1% | 1.9% | 0% | 0% | 0% |
| Technical issues occurred during the VR-training | 46.2% | 40.4% | 7.7% | 1.9% | 3.8% |

tion workers on construction sites do not involve sitting indoors, which was considered tiring.

### 5.5.2. Process analysis of HFST training

HFST training participants evaluated the extent (1 = not at all, 7 = very much) to which the training had increased their knowledge regarding factors affecting the occurrence of accidents ($M$ = 5.44, SD = 1.20, min = 3 max = 7) and to which the training had increased their safety knowledge ($M$ = 4.67, SD = 1.7, min = 2 max = 7). The study participants also rated the extent to which the contents of the HFST training were difficult to understand ($M$ = 2.40, SD = 1.38, min = 1, max = 7) and the extent to which using the HFST viewpoint was easy ($M$ = 5.06, SD = 1.17, min = 1 max = 7). In general, these results indicated that study participants perceived HFST training positively. We found that the engagement of study participants in HFST training activities was higher than that of those in the lecture-based safety training. However, the level of discussion between the study participants during the HFST training varied across the study organizations (11.1% no discussion during training, 33.1% some discussion, 11.1% a moderate amount of discussion, 33.3% quite a lot of discussion, 11.1% a great deal of discussion). Safety knowledge sharing activity varied among the study participants during the HFST training (22.2% none, 44.4% some safety knowledge sharing, 11.1% a moderate amount of knowledge sharing, 22.2% quite a lot of knowledge sharing). Similarly, the amount of comments made by the study participants varied across the HFST training (44.4% some comments, 11.1% a moderate amount of comments during training, 33.3% quite a lot of comments during training, 11.1% a great deal of comments during training). Of the employees who participated in the HFST training and also responded to the telephone interviews ($n$ = 47), 82% gave positive feedback on the training. According to the employees, the training led to a more in-depth consideration of the issues affecting occupational safety and the underlying causes of accidents. Sixteen percent of the participants also gave negative feedback about the training, mainly related to the fact that some felt it did not give them any new knowledge.

## 6. Discussion

The process analysis results showed that all the intervention processes were mostly delivered as intended. There were minor technical discrepancies in the VR-based safety training and some variation in terms of active participation during the HFST training. The efficacy evaluation showed that the VR-based safety training was more effective than the lecture-based safety training, especially in enhancing safety self-efficacy in identifying factors affecting safety. This result is somewhat similar to the results of previous research showing that VR-based training strengthens hazard recognition skills in particular (Perlman, Sacks, & Barak, 2014). According to Bandura (1997), mastery experience is the most important source of self-efficacy. In our study, verbal persuasion was the key mechanism for enhancing self-efficacy in the lecture-based safety training, while the VR-based safety training also provided mastery experiences of safety-related actions during the tasks in the virtual environment. Therefore, the VR-based training offered a more effective method for boosting safety self-efficacy. Our results also indicated that the VR-based safety training's effects on self-efficacy in identifying factors affecting safety were sustained in the one-month follow-up measurements. According to our results, the VR-based safety training had a stronger impact on outcome expectancies related to worker-level safety performance at short-term follow-up. Outcome expectancies were targeted in the lecture-based safety training by providing information about the relationship between safety-related actions and accident prevention. In the VR-based training, the study participants also removed hazards in a virtual environment. Thus, informational influence was supplemented by actual observation of the effects of one's own actions. More generally, the participants perceived VR-based training as more inspiring, which supported our finding that VR-based training had a stronger impact on safety motivation.

We also evaluated the efficacy of a participatory HFST training program. Contrary to our hypothesis, no positive effects emerged in comparison to the control group without HFST training. Several factors may have contributed to this result. First, due to our limited sample size, our analyses may have been underpowered for detection of a smaller intervention impact. Furthermore, observational data indicated that the level of participants' activity (engagement) varied across the study organizations, which may have affected to the results concerning efficacy of HFST training. Previously, Teperi and colleagues (2015) found positive effects of human factors training, however, they implemented human factors safety training on both managerial and employee levels. In our study, training targeted only the worker level. This in turn may have led to the participants feeling powerless or sidelined, with no support or sufficient understanding from supervisory and managerial levels. Moreover, in the previous studies (Teperi et al., 2015, 2018), the human factor safety training was more extensive than that in our study and included substantially longer training sessions and workshops lasting from one to several days, whereas in our study, the HFST training lasted only around 1.5 h. Therefore, it is possible that a more extensive HFST training would be needed in order to identify effects on the measures used in this study. Although we could not identify any statistically significant differences between the increase of safety self-efficacies or outcome expectancies of the study participants with HFST training and the control group without HFST training, the qualitative data do indicate some positive effects. Specifically, the study participants perceived that HFST training increased in-depth consideration of the issues affecting occupational safety and the underlying causes of accidents. However, we did not measure these perceptions in our outcome measures, and this should be examined in future studies.

## 7. Study limitations

Our study has a few limitations. Due to the low reliability of the measure of internal safety locus of control, the effects can be considered only indicative. It is also important to acknowledge that the majority of the detected effect sizes were small – a likely reason being the relatively high baseline levels of participants' safety competencies. Nevertheless, even small intervention effects can translate into substantial safety outcomes. This is illustrated by our interview data, which provided examples of workplace activities after the training.

We also wish to underline that the lecture-based safety training in our study was implemented using a trainer-centered approach. This should be considered when interpreting the results. However, this is an informative point of reference, because safety training in the workplace is often based on a passive approach (see Burke et al., 2006). Furthermore, self-reported safety performance measures should be interpreted as indicative and not definitive evidence of behavioral outcomes. Finally, our sample size was rather small, which was also reflected in the confidence intervals of the results.

## 8. Study strengths

To our knowledge, this is one of the few RCTs to investigate the impact of VR-based safety training and was the first study to explore the efficacy of HFST training in the construction sector. We also complemented the efficacy assessment of intervention processes with process analysis, which is considered important in safety intervention studies. Process analysis provides valuable information for both practical training planning and new research topics.

## 9. Practical implications

Our study showed that virtual reality is a potential tool in safety training for increasing safety competencies and fostering motivational change in terms of the safety performance of construction sector workers. To avoid any misinterpretation, we do not propose that virtual reality could completely replace other safety training methods (e.g., hands-on-training, safety discussions). Our study indicates that VR provides a promising alternative to passive learning methods. VR safety training provides opportunities for active learning and its motivating effect complements more traditional training activities (e.g., lectures). Our process analysis revealed that even small technical issues (e.g., positional tracking issues) may influence the learning experiences and outcomes related to VR-based training. We had direct control over our first-party developed software solution (the VR platform software and the created VR content running on it), and there were no apparent software bugs in the developed platform or the training content during the intervention. No software solution could have resolved the technical issues that arose concerning the hardware, which was provided by a third party. This covers both the VR headset and the computer communicating with it, and the interplay between the VR headset and a computer can be prone to incompatibility issues due to their different providers. However, this highlights the importance of pilot studies and pre-testing technology-based safety training methods before proceeding to RCTs.

Producing immersive VR environments can be time consuming and requires both technical and substance-driven competence for providing a successful learning experience. The production of VR content can be facilitated with a proper content creation pipeline, and a dedicated software framework. This study utilized a dedicated, domain-specific approach that was first created and then used to model and implement all the training scenarios for the project (Nykänen et al., 2020). Having similar tools at hand can support dissemination and implementation of VR-safety training methods at workplaces.

Our qualitative data indicates that HFST training may have positive effects on safety perceptions. In the future, HFST training should be implemented on both managerial and employee levels in order to enable a shared understanding of human factors affecting safety throughout the organization. Furthermore, HFST training should be given more time to provide ample space for discussions and practice, hence enhancing the learning results.

## 10. Future studies

Many important aspects of the present study need further research: First, using larger sample sizes, future studies should compare the outcomes of immersive VR-based safety training and more participatory face-to-face safety training. Previous studies have also pointed out that implementation process factors are potential moderators of intervention outcomes (see Carroll et al., 2007). The overall effectiveness of virtual environments may be dependent on a multitude of factors such as gamification, usability, and virtual environment design. The success of the resulting VR training thus depends on more factors than merely the immersive capabilities provided by the hardware used. The immersive capabilities supported by the technology can easily be exploited for ill-fitting user experience, if used without taking the user experience and engagement into account. For creating a virtual reality training environment suitable for engaging self-guided learning, all the contributing factors present in interactive content such as VR should be considered. Future studies should explore the associations between key characteristics of the virtual learning experience (e.g., sense of presence and usability). This can provide insight over how an effective VR training should fundamentally be built and to what extent should immersion be depended on for explaining these achieved results. Studying these aspects could increase understanding of the key "active ingredients" of VR-based safety training (see Abry, Hulleman, & Rimm-Kaufman, 2015). Furthermore, due to our limited study sample, it was not possible to further test the moderating role of technical issues during the VR-based safety training and intervention outcomes. Future studies should assess the impact of possible technical issues on learning outcomes.

Future studies should also explore the role of trainees' activity levels in HFST training outcomes and evaluate the effects of HFST training on trainees' perceptions of the causal explanations and attributions of workplace accidents. Furthermore, it would be important to understand how the broader safety knowledge provided by the HFST training can be transformed to better safety practices on individual, team, and organizational levels.

## Acknowledgements

This article is based on research supported by the Finnish Work Environment Fund (#117306). We thank Alice Lehtinen for the language editing.

## Declaration of interest

The authors declare that they have no conflicts of interest.

## References

Abry, T., Hulleman, C. S., & Rimm-Kaufman, S. E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation, 36*(3), 320–338.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and human Decision Process, 50*, 179–211.

Nykänen, M., Puro, V., Tiikkaja, M., Kannisto, H., Lantto, E., Simpura, F., Uusitalo, J., Lukander, K., Räsänen, T., & Teperi, A.-M. (2020). Evaluation of the efficacy of a virtual reality-based safety training and human factors training method: study protocol for a randomised-controlled trial. *Injury Prevention, 26*, 360–369.

Arezes, P., & Miguel, A. (2008). Risk perception and safety behaviour: A study in an occupational environment. *Safety Science, 46*, 900–907.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review, 84*, 191–215.

Bandura, A. (1997). *Self-efficacy. The exercise of control*. NY: W.H Freeman and Company.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9–44.

Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science, 12*(1), 23–33.

Bouchard, S., Robillard, G., & Renaud, P. (2007). Revising the factor structure of the simulator sickness questionnaire. *Annual Review of Cybertheraby and Telemedicine, 5*, 117–122.

Burke, M. J., Salvador, R., Smith-Crowe, K., Chan-Serafin, S., Smith, A., & Sonesh, S. (2011). The dread factor: How workplace hazards and training influence learning and performance. *Journal of Applied Psychology, 96*, 46–70.

Burke, M. J., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. O., & Islam, G. (2006). Relative effectiveness of worker safety and health training methods. *American Journal of Public Health, 96*, 315–324.

Burke, M., Scheuer, M., & Meredith, R. (2007). A dialogical approach to skill development: The case of safety skills. *Human Resource Development Journal, 17*, 235–250.

Carroll, C., Patterson, M., Wood, S., Booth, A., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*, 40.

Casey, T., Krauss, A., & Turner, N. (2018). The one that got away: Lessons learned from the evaluation of a safety training intervention in the Australian prawn fishing industry. *Safety Science, 108*, 218–224.

Christian, M., Bradley, J., Wallace, J., & Burke, J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology, 94*, 1103–1127.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
Curcuruto, M., Mearns, K., & Mariani, M. (2016). Proactive role-orientation toward workplace safety: Psychological dimensions, nomological network and external validity. *Safety Science, 87*, 144–155.
Dejoy, D. (1996). Theoretical models of health behavior and workplace self-protective behavior. *Journal of Safety Research, 2*, 61–72.
Douglas, D., & Peucker, T. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization, 10*, 112–122.
Dugan, A. G., & Punnett, L. (2017). Dissemination and implementation research for occupational safety and health. *Occupational Health Science, 1*, 29–45.
Dużmańska, N., Strojny, P., & Strojny, A. (2018). Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. *Frontiers in Psychology, 9*.
Eisinga, R., Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? *International Journal of Public Health., 58*, 637–642.
Ford, M. T., & Tetrick, L. E. (2008). Safety motivation and human resource management in North America. *The International Journal of Human Resource, 19*, 1472–1485.
Gao, Y., Gonzalez, V., & Yiu, T. (2019). The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review. *Computers & Education, 138*, 101–115.
Guo, H., Li, H., Chan, G., & Skitmore, M. (2012). Using game technologies to improve the safety of construction plant operations. *Accident Analysis and Prevention, 48*, 204–213.
Gheisari, M.; Esmaeli, B. (2019). CPWR Small Study Program (report number 17-7-PS) - PARS: Using augmented panoramas of reality for construction safety training.
Glendon, A. I., Clarke, S., & McKenna, E. F. (2006). *Human safety and risk management*. Boca Raton, FL: CRC Press, Taylor & Francis.
Grabowski, A., & Jankowski, J. (2015). Virtual reality-based pilot training for underground coal miners. *Safety Science, 72*, 310–314.
Griffin, M., & Neal, M. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*, 347–358.
Guerin, R. J., Toland, M. D., Okun, A. H., Rojas-Guyler, L., Baker, D. S., & Bernard, A. L. (2019). Using a modified theory of planned behavior to examine teachers' intention to implement a work safety and health curriculum. *Journal of School Health, 89*, 549–559.
Hartmann, T., Wirth, W., Vorderer, P., Klimmt, C., Schramm, H., & Böcking, S. (2015). Spatial presence theory: State of the art and challenges ahead. In M. Lombard, F. Biocca, J. Freeman, W. Ijsselsteijn, & R. Schaevitz (Eds.), *Immersed in Media*. Cham: Springer.
Jones, J., & Wuebker, L. (1993). Safety locus of control and employees accidents. *Journal of Business and Psychology, 7*, 449–457.
Katz-Navon, T., Naveh, E., & Stern, Z. (2007). Safety self-efficacy and safety performance. *International Journal of Health Care Quality Assurance, 20*, 572–584.
Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3*(3), 203–220.
Lawson, B. D., Graeber, D. A., Mead, A. M., & Muth, E. R. (2002). Signs and symptoms of human syndromes associated with synthetic experiences. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 589–618). Mahwah, NJ: Lawrence Erlbaum Associates Inc..
Leder, J., Horlitz, T., Puschmann, P., Wittstock, V., & Schütz, A. (2019). Comparing immersive virtual reality and powerpoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making. *Safety Science, 111*, 271–286.
Lee, E., Wong, A., & Fung, C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education, 55*, 1424–1442.
Li, X., Yi, W., Chia, H.-L., Wang, X., & Chan, A. (2018). A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation in Construction, 86*, 150–162.
Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence, journal of computer-mediated. *Communication, 3*.
Mantovani, F., & Castelnuovo, G. (2003). The sense of presence in virtual training: enhancing skills acquisition and transfer of knowledge through learning experience in virtual environments. In G. Riva & F. Davide (Eds.), *Being there: Concepts, effects and measurement of user presence in synthetic environments* (pp. 167–182). Amsterdam: IOS Press.
Mazaheri, M., Hidarnia, A., & Ghofranipour, G. (2012). Safety education and control: A tool to measure the safety locus of control. *Journal of Education and Health promotion, 1*.
McCreery, M., Schrader, P. G., Krach, S. K., & Boone, R. (2013). A sense of self: The role of presence in virtual environments. *Computers in Human Behavior, 29*, 1635–1640.
Merta, M., & Kelling, N. (2015). Effects of level of control on simulator sickness using a virtual reality head mounted device. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59*(1), 766–769.
Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*, 364–386.
Neal, A., & Griffin, M. (2006). Study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*, 946–953.
Neal, A., Griffin, M., & Hart, P. (2000). The impact of organizational climate on safety climate and individual behavior. *Safety Science, 34*, 99–109.
Nykänen, M., Salmela-Aro, K., Tolvanen, A., & Vuori, J. (2019). Safety self-efficacy and internal locus of control as mediators of safety motivation – Randomized controlled trial (RCT) study. *Safety Science, 117*, 330–338.
Nykänen, M., Sund, R., & Vuori, J. (2018). Enhancing safety competencies of young adults. A randomized field trial (RCT). *Journal of Safety Research, 67*, 45–56.
Perlman, A., Sacks, R., & Barak, R. (2014). Hazard recognition and risk perception in construction. *Safety Science, 64*, 22–31.
Pinto, A., Nunes, I. L., & Ribeiro, R. A. (2011). Occupational risk assessment in construction industry – Overview and reflection. *Safety Science, 49*(5), 616–624.
Plechatá, A., Sahula, V., Fayette, D., & Fajnerová, I. (2019). Age-related differences with immersive and non-immersive virtual reality in memory assessment. *Frontiers in Psychology, 10*, 1330.
Osborne, J. W., & Overbay, A. (2004). *The power of outliers (and why researchers should always check for them)*. Practical Assessment, Research & Evaluation.
Ramer, U. (1972). An iterative procedure for the polygonal approximation of plane curves. *Computer graphics and Image Processing, 1*, 244–256.
Ricci, F., Chiesi, A., Bisio, C., & Pelosi, A. (2016). Effectiveness of occupational health and safety training. A systematic review with meta-analysis. *Journal of Workplace Learning, 28*, 355–377.
Robson, L., Stephenson, C., Schulte, P., et al. (2012). A systematic review of the effectiveness of occupational health and safety training. *Scandinavian Journal Work Environmental Health, 38*, 193–208.
Rotter, J. (1966). Generalized expectancies for internal versus external locus of control. *Psychology Monographs, 80*.
Rotter, J. (1982). *The development and applications of social learning theory*. New York: Praeger.
Sacks, R., Perlman, A., & Barak, R. (2013). Construction safety training using immersive virtual reality. *Construction Management and Economics, 31*, 1005–1017.
Salzman, M., Dede, C., Loftin, R., & Chen, J. (1999). A model for understanding how virtual reality aids complex conceptual leaning. *Presence: Teleoperators and Virtual Environments, 8*, 293–316.
Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence, 10*, 266–281.
Schulz, K., Altman, D., & Moher, D. (2010). Moher D. Consort 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*.
Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence-Teleoperators and Virtual Environments, 6*, 603–616.
Teperi, A.-M., Leppänen, A., & Norros, L. (2015). Application of new human factors tool in an air traffic management organization. *Safety Science, 73*, 23–33.
Teperi, A.-M., Puro, V., Tiikkaja, M., et al. (2018). *Developing and implementing a human factors (HF) tool to improve safety management in the nuclear industry* (p. 2018). Helsinki: Finnish Institute of Occupational Health.
Teperi, A.-M., Puro, V., & Ratilainen, H. (2017). Applying a new human factor tool in the nuclear energy industry. *Safety Science, 95*, 125–139.
van der Molen, H., Basnet, P., Hoonakker, P., Lehtola, M., Lappalainen, J., Frings-Dresen, M., et al. (2018). Interventions to prevent injuries in construction workers. *Cochrane Database Systematic Review, 48*.
Witmer, B., Jerome, C., & Singer, M. (2005). The factor structure of the presence questionnaire. *Presence, 14*, 298–312.
Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*, 225–240.
Wójcicki, T., White, S., & McAuley, E. (2009). Assessing outcome expectations in older adults: The Multidimensional Outcome Expectations for Exercise Scale (MOEES). *The Journals of Gerontology, Series B: Psychological Science and Social Sciences, 64*, 33–40.
Zhou, Z.; Goh, Y.; & Li, Q. (2015). Overview and analysis of safety management studies in the construction industry, 72, 337–350.

**Mikko Nykänen** (Msc. researcher, PhD.student) works at the Finnish Institute of Occupational Health. He is actively involved in the safety training and safety intervention projects. The common themes running through his work include developing behavior change interventions, planning RCT field trials and supporting school-to-work transition of young workers.

**Maria Tiikkaja**, PhD, Specialist Research Scientist, at Finnish Institute of Occupational Health. Research work and interest include occupational safety and applicability of new technologies as well as safety culture. Chair of The Ethical Committee which is an institutional review board

**Eero Lantto** (M.Soc.Sci, researcher) works at the Finnish Institute of Occupational Health. He is a professional in developing safety and health through research and development projects for different industries.

**Henriikka Kannisto** (MSc (tech.), research engineer) works both with occupational safety as well as safety critical fields as a researcher, developer and trainer at the Finnish Institute of Occupational Health. Kannisto has her background in engineering and safety management, and during the last 9 years she has extended her knowledge to cognitive psychology in safety. She has expertise in the fields of human factors, safety values and safety culture.

**Tuula Räsänen** (Lic.Soc.Sc., PhD (Technology), Senior Specialist) works at the Finnish Institute of Occupational Health. She has been involved in safety research, safety expertise work and safety training at workplaces over thirty years. Her current work involves studying of new VR-methods for training of occupational safety, especially for construction sites and studying of silent signals of big data material concerning to occupational accidents in Finland.

**Vuokko Puro** (M.Sc. and PhD.student) works at the Finnish Institute of Occupational Health as safety specialist. She is actively involved in the occupational safety and human factors projects. The common themes running through her work include new technologies in occupational safety and developing better mastery of human factors at safety critical fields.

**Dr. Anna-Maria Teperi** is Chief Researcher in the Finnish Institute of Occupational Health (FIOH). She is an Adjunct Professor at Tampere University in the field of Human Factors in Safety Management, with 20 years of experience in research and implementing HF in safety-critical fields, e.g. aviation, the nuclear industry, maritime and railway.

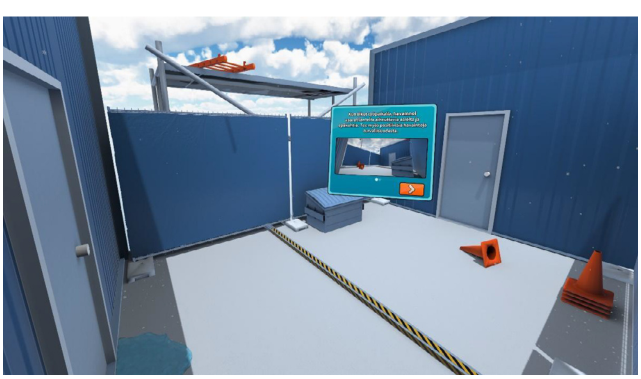
**Tarja Heikkilä** (Senior Specialist, MSc Tech.) works at the Finnish Institute of Occupational Health. She is involved in the occupational safety, risk assessment and safety training in construction. The common themes running through her work include developing processes to promote occupational safety.

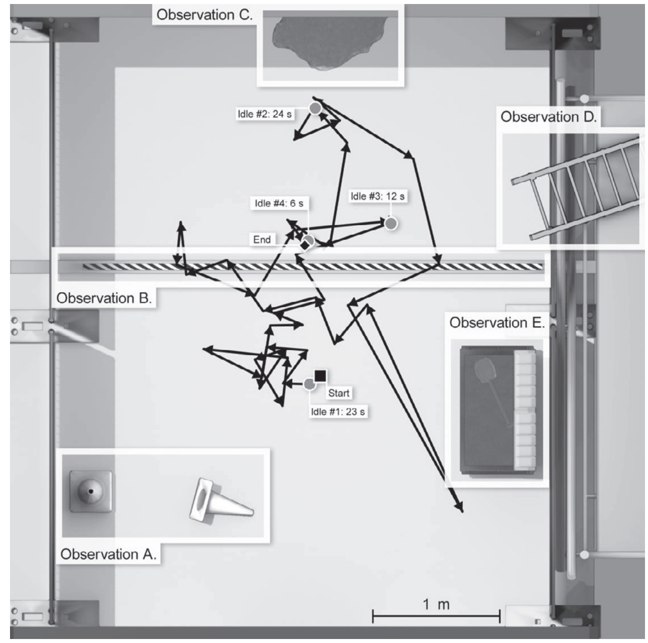
**Jose Uusitalo** (B. Eng specialist) works at the Finnish Institute of Occupational Health. He has worked as a software engineer, a 3D artist, and a photographer in projects exploring the use of digital tools for occupational health services and developing occupational safety training solutions for web browsers and for virtual reality platforms. His current work focuses on multimedia content production and developing internal tools for content production and management.

**Frans Simpura** (MSc in Computer Science) works at the Finnish Institute of Occupational Health as a Senior Specialist. He is the lead developer of the virtual reality implementation of the research project and the virtual reality learning platform Virtuario™. He is devoted in finding novel solutions for developing easy-to-use and captivating virtual reality applications.

**Kristian Lukander** (M.Sc.(Tech), Senior Specialist, PhD student) is actively involved in psychophysical research, VR development, and product development at the Finnish Institute of Occupational Health. His research interests include methods for measuring psychophysiological status, and the more cognitive aspects of work.

## Appendix A






The figure in Appendix A provides an illustrative example of active learning activities during a VR-training. The example describes a sample walking path in one scene of one of the trainings. The user started this training scene in the square labeled "Start" and ended in the diamond labeled "End". The user's movement path between these points has been simplified using the Ramer–Douglas–Peucker algorithm (Ramer, 1972; Douglas & Peucker, 1973) into several movement segments illustrated with arrows to preserve directional information, however, the length of the arrow does not imply faster movement speed. Gray circles labeled "Idle #1" through "Idle #4" show the locations in the scene where the user was mostly still, how many seconds the user was still for, and the order in which these events took place. White frames labeled "Observation A." through "Observation E." surround all interactive parts of the scene that have educational information attached. When activating one of these observations the educational information is displayed in a text box near the object.

| Label | Object | Observation |
|---|---|---|
| A. | Loose traffic cones | A tripping hazard: keep walkways clear of objects |
| B. | Cable channel | A potential tripping hazard: marked with caution tape |
| C. | Ice | A slipping hazard: be aware of locations where ice may form |
| D. | An extension ladder on scaffolds | A falling object hazard: objects should be stored in stable locations |
| E. | A gravel container | The user activates the container to fix the hazard at Observation C |

## Appendix B

| Learning topic | Learning topics in HFST training |
|---|---|
| 1. | **Individual level**<br>Identifying how personal factors such as mastery of work and, psychological and attitudinal factors are connected to work-related accidents |
| 2. | **Work level**<br>Identifying how the work procedures and work distribution contribute to or prevent accidents |
| 3. | **Group level**<br>Identifying how the structure or cohesion of working groups and information flow influence occupational safety processes |
| 4. | **Organizational level**<br>Identifying how decisions concerning resources and co-operation between different organizational levels are related to occupational safety |
| 5. | **Contractor and network level**<br>Identifying how inter-actor communication, collaboration and roles are related to occupational safety |

## Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsr.2020.09.015.