buildings

Article

A Task-Centred Methodology to Evaluate the Design of Virtual Reality User Interactions: A Case Study on Hazard Identification

Pierre Raimbaud [1,2,*,†], Ruding Lou [1], Florence Danglade [1], Pablo Figueroa [2], Jose Tiberio Hernandez [2] and Frederic Merienne [1]

[1] Arts et Metiers Institute of Technology, LISPEN, HESAM Université, 71100 Chalon-Sur-Saone, France; ruding.lou@ensam.eu (R.L.); florence.danglade@ensam.eu (F.D.); frederic.merienne@ensam.eu (F.M.)

[2] Systems and Computing Engineering, Universidad de los Andes, Bogota 111711, Colombia; pfiguero@uniandes.edu.co (P.F.); jhernand@uniandes.edu.co (J.T.H.)

* Correspondence: pierre.raimbaud.pr@gmail.com or p.raimbaud@uniandes.edu.co

† Current address: Inria Rennes-Bretagne Atlantique, Campus de Beaulieu, 263 Avenue Général Leclerc, 35042 Rennes, France.

Abstract: Virtual reality (VR) is a computer-based technology that can be used by professionals of many different fields to simulate an environment with a high feeling of presence and immersion. Nonetheless, one main issue when designing such environments is to provide user interactions that are adapted to the tasks performed by the users. Thus, we propose here a task-centred methodology to design and evaluate these user interactions. Our methodology allows for the determination of user interaction designs based on previous VR studies, and for user evaluations based on a task-related computation of usability. Here, we applied it on the hazard identification case study, since VR can be used in a preventive approach to improve worksite safety. Once this task and its related user interactions were analysed with our methodology, we obtained two possible designs of interaction techniques for the worksite exploration subtask. About their usability evaluation, we proposed in this study to compare our task-centred evaluation approach to a non-task-centred one. Our hypothesis was that our approach could lead to different interpretations of user study results than a non-task-centred one. Our results confirmed our hypothesis by comparing weighted usability scores from our task-centred approach to unweighted ones for our two interaction techniques.

Keywords: construction safety; hazard identification; virtual reality; user interactions; user-centred design; task-centred design

1. Introduction

The Architecture, Engineering and Construction (AEC) industry is a sector where the design and the preparation of future elements and actions are key to the whole lifecycle of a building [1]. This goes from its preconception with ground and cost studies to its maintenance with future reparation planning. To address such needs of foresight, anticipation, and preventive correction, this sector is currently using more and more computer technologies [2,3]. Among them, building information modelling (BIM) methodology and tools are ones of the most widely used, notably for their power of data follow-up, update, and sharing [4,5]. Nonetheless, other computer technologies can also be used in the industry such as virtual reality (VR), augmented reality (AR), the internet of things, or artificial intelligence [6]. Indeed, for example, the VR technology has been particularly adopted during the design stage of the construction [7], and AR during the maintenance stage [8].

In the last years, it has been shown in the literature that computer-based technologies can be beneficial for a large variety of purposes and issues in the AEC industry, such as energy management [9], construction monitoring [10], or design reviewing [7,11].

Focusing here on one cross-cutting issue in particular, safety is a paramount one in the AEC industry, both before and after the construction of a building. Indeed, on one hand, during the operation and maintenance stages, i.e., after the construction, the safety of the building users is required and can be ensured by carrying out corrective and follow-up actions through safety assessments [12]. To conduct such safety assessments, computational approaches such as artificial neural networks can be used and contribute to the prediction of potential safety problems and, thus, to the anticipation of maintenance actions, as explained by Harirchian et al. [13]. On the other hand, before the use of the building and, therefore, during its construction, this is the safety of the AEC workers that must be ensured. The AEC industry is indeed an industry that is particularly dangerous for its workers [14–17]. In many countries, accidents are more frequent and more fatal for workers in the AEC industry than in other ones [18]. To address this crucial issue [14], computer-based approaches such as BIM-based ones have been proposed in the literature [19,20].

In this paper, for our case study, we focused on the construction workers' safety, on a worksite. According to the Occupational Safety and Health Administration—the OSHA—the most common construction hazards are the fall, electrocution, caught in between, and struck-by hazards [21]. To address these safety issues, one option is to conduct preventive and corrective actions on the worksites during the construction, for example using signals that highlight hazardous zones in augmented reality [22] or using monitoring systems that can prevent mechanical failures [23]. Another option is to take safety measures before the beginning of the construction, notably by improving the understanding of these worksite hazards by the AEC professionals [24], and by learning more about the ways to identify and mitigate them [25,26]. As a result, a technique called design for safety has been proposed in the literature to prevent hazards from the design stage of the building lifecycle [27–29]. This has been particularly studied in BIM authoring tools, either to avoid fall hazards with the automatic addition of protections in such hazardous zones [20,30], or to prevent electrocution fatalities [31]. For the two other main kinds of hazards—caught between and struck-by ones—other preventive actions can be done on construction planning, instead of the building design itself. Indeed, these kinds of hazards are directly related to the dynamic aspect of a worksite since they are mostly due to construction vehicles and workers movements [32,33]. Then, as construction vehicles and workers follow planned movements during the construction, hazardous zones—zones with conflicts that may result in collisions—can be identified in construction planning, usually using paper plans [34] or eventually BIM 4D simulations [35], i.e., construction simulation with 3D representations over time.

In our approach to improve the workers' safety, we focused on the preventive actions aforementioned, and in particular on the process of conducting hazard identification reviews. For that, the VR technology can be used [36], notably for hazards that are related to dynamic events, thanks to the improved sense of immersion and presence that VR offers [37–39]. In that sense, several studies have been conducted in the literature about the use of VR for hazard assessment [40]. Many of them have focused on hazard identification in VR for training purposes, either for workers or civil engineers [41]. Indeed, Sacks et al. [42] compared the effect of a traditional training to a VR training on the civil engineers' memory by evaluating their hazard knowledge after different intervals of times. In their experiment, the VR-trained participants performed better than the control group who used paper plans. Zhao et al. [43,44] and Fang et al. [45] focused their studies on training workers to avoid having hazardous behaviours, for example when they use a crane. Finally, in the same line, Xu et al. [46] and Joshi et al. [47] conducted objective and subjective user evaluations on VR training for safety, which confirmed the benefits of this technology for construction workers. Nonetheless, the use of VR has also been studied for other purposes than training, and notably as an environment that can allow safety managers to detect design or construction planning defects that would engender hazardous zones in the future worksite [48–51].

In line with the aforementioned papers, which showed the benefits of VR for the hazard identification task, our paper proposes to study the use of VR environments for this specific task, and its design and evaluation process in particular. Indeed, as virtual reality is intrinsically linked to human factors [52,53], one issue related to VR is that the design of VR environments and notably its user interactions should be adapted to the VR users' tasks [54]. Indeed, the design of the user interactions in VR is paramount since their usability is directly linked to user task performance [55,56], and as a result to the effectiveness of the user tasks performed in VR [57]. To address this issue, user-centred design methodologies have been developed in the VR literature. These methodologies classified the different existing interaction techniques and presented some approaches to make design choices for the VR user interactions [58–60]. Nonetheless, a remaining problem is that in these methodologies the main actors of the design process must have knowledge expertise in VR [58]. Indeed, such user-centred methodologies appear to be conceived to be used by VR developers and then to guide these developers in a process of integration of the users [60]. As a result, the influence of the users and also of the user tasks characteristics on the design may be lower than expected. To address this problem, we propose here a new methodology for the design and the evaluation of VR user interactions, which is centred on the tasks performed by the users and makes them the main actors of the VR design process. The novelty of our work lies in this task-centred approach for both the design and the evaluation of VR user interactions.

About our task-centred methodology and our evaluation approach in particular, we first followed the principles stated by Bowman et al. about usability [61,62]. Therefore, we mainly based our evaluation approach on the use of quality factors to define and determine the usability of VR user interaction techniques. Then, we aimed to strengthen the use of these usability quality factors in a task-centred way. Hence, we proposed to attribute different weights to the quality factors, which represent their relative importance for the current task. Then, these weights could be used to perform a task-centred computation of the usability of an interaction technique. In the user evaluation that we conducted in this paper, we computed both weighted-sum and unweighted (simple-sum) usability scores, to compare our task-centred evaluation approach with a non-task-centred one [59,60]. We expected that weighted and unweighted scores would give different interpretations about the usability of two different interaction techniques that we tested here. We present in Section 2 our task-centred methodology, how we applied it on the hazard identification case study, and the user evaluation that we conducted to test our hypothesis. Results of our experiment are given in Section 3. Finally, Section 4 provides an interpretation of these results, and Section 5 gives some conclusions regarding this study and ideas of future work.
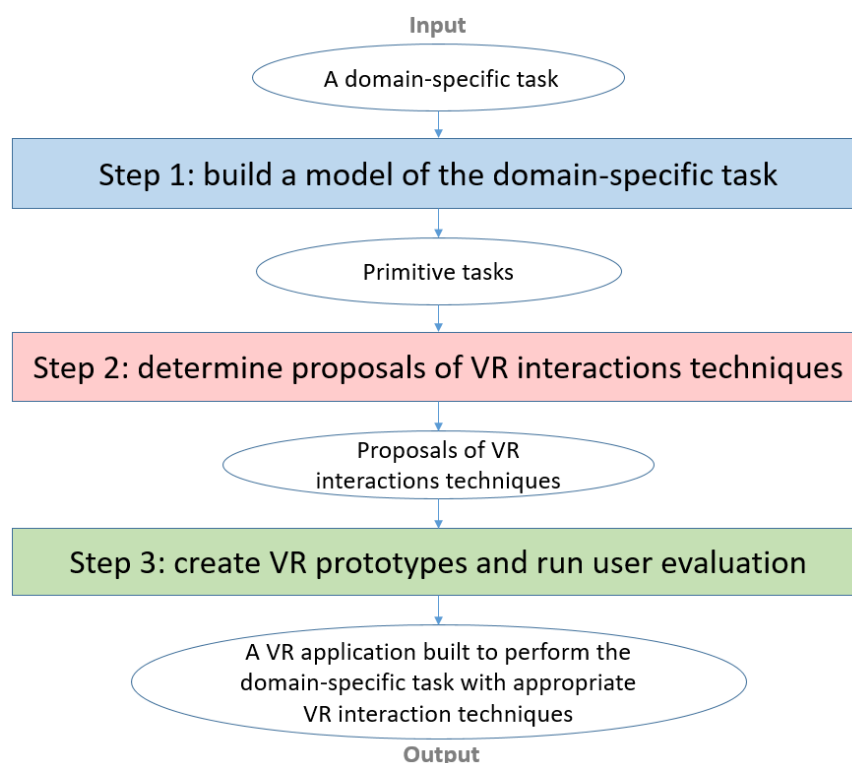
2. Materials and Methods

2.1. Our Task-Centred Methodology to Design and Evaluate VR User Interactions

In our research, we proposed a new task-centred methodology to design and evaluate VR user interactions. Its objective is to guide the creation of VR applications and in particular to drive the choices of VR interactions in terms of usability according to the user task. Our methodology relies on taxonomies and characterisations of tasks, user interactions, and user interaction techniques, and is supported by previous works in this direction from the VR literature [58,60,63]. This methodology is composed by three task-centred steps, two to theoretically design the VR user interactions and one to proceed to a concrete user evaluation. Figure 1 gives an overview of our methodology.

The first step of our methodology consists of the construction of a model that defines and decomposes the user task to be performed. In that purpose, a hierarchical task analysis [64] is conducted on this task, which leads as an output to a decomposition into subtasks that we called primitive tasks. The aim of such analysis is to go from a task that is specific to a field to more generic primitive tasks that can be directly linked to VR user interactions, such as selection, manipulation, and navigation. Examples of such primitive

subtasks are "to move to some places", or "to orient an object". To perform this hierarchical task analysis, we created a semi-automated system that guides the task decomposition into subtasks.

Input

A domain-specific task

Step 1: build a model of the domain-specific task

Primitive tasks

Step 2: determine proposals of VR interactions techniques

Proposals of VR interactions techniques

Step 3: create VR prototypes and run user evaluation

A VR application built to perform the domain-specific task with appropriate VR interaction techniques

Output

Figure 1. Overview of our task-centred methodology to design and evaluate VR user interactions.

Then, our second step consists of the determination of proposals of interaction techniques that are adapted to the subtasks obtained from the first step. This step uses a rule-based system to deduce the characteristics of the interaction technique that would be more adapted in terms of usability for a given subtask, according to its characteristics. This system relies on usability studies that have been conducted in the literature for different kinds of tasks and conditions using different VR interaction techniques. Based on these studies and depending on the kind of interaction—selection, manipulation, or navigation— and the user task characteristics, our system can determine either a unique proposal of interaction technique, or several proposals in absence of previous results for some VR interaction technique characteristics. Examples of such interaction technique characteristic are the order of magnitude of the speed for a navigation interaction, or the kind of selection tool for a selection interaction.

Finally, our methodology ends with an evaluation step that consists of the creation of a VR application with the proposed interaction techniques and of a task-centred user study using this VR application. Each primitive task—determined in our step 1—and its associated interaction technique(s)—determined in our step 2—should be evaluated independently in the user study. Moreover, in the case of multiple proposals of interaction techniques for one user interaction, these different proposals must be tested in the user study, and their usability can be compared to improve the knowledge about such VR interaction techniques and their use for the studied user task. In any cases, first, VR applications must be developed as concrete supports for the user study. At this stage, external constraints that may influence the concrete implementation of an interaction technique, such as the budget or the kind of devices to be used, must be stated and considered. Then, a task-centred user study must be conducted, based on one main principle: the application of quality factors on the usability evaluation [61,65], according to the studied subtask. The quality factors of an interaction vary depending on the kind of interaction, and can

be for example the spatial awareness for a navigation interaction, or the ease-of-use for a selection interaction. Our task-centred approach relies on defining the relative importance of each quality factor depending on the current user task and subtask. This can notably be done thanks to the users' expertise in the field to which the task belongs. To define such relative importance, we propose in our methodology to use weights that represent the importance of each quality factor in terms of usability for the VR interaction when performing the current user task. Finally, a weighted usability score can be computed using a formula that takes in account the user scores for each quality factor and the weights previously determined. Additionally, to reduce the number of quality factors and, thus, the number of measures to take, a preliminary selection of the quality factors that are relevant for the user task can be done first, through an interview with professionals of the task field.

2.2. Objective of the Study and Hypothesis

In this paper, we present how we applied our task-centred methodology on the case study of the hazard identification task. This methodology allowed us to determine the design of the VR user interactions for each subtask; nonetheless, for the worksite exploration subtask, we determined two possible designs for its associated navigation user interaction. These two designs can be evaluated and compared in terms of usability in the evaluation step of our task-centred methodology. In this study, we propose to follow both our task-centred evaluation approach and a non-task-centred one. Our objective is to study the effect of the application of our approach on the usability results and their interpretation, compared to a non-task-centred approach.

In that purpose, we propose here to compute the usability scores for our two possible interaction techniques in two main different ways: an unweighted way from a non-task-centred approach, and the weighted way from our methodology. Figure 2 shows these two computational approaches and their associated formulas, U1 in the unweighted way and U2 in the weighted way [66]. Then, we propose to compare the unweighted and weighted scores for our two navigation interaction techniques, and to see if these comparisons are similar or different depending on the kind of computation. Our **hypothesis** is that our **task-centred evaluation** with the weighted computation would give **different results** and usability interpretation compared to the **non-task-centred** evaluation.
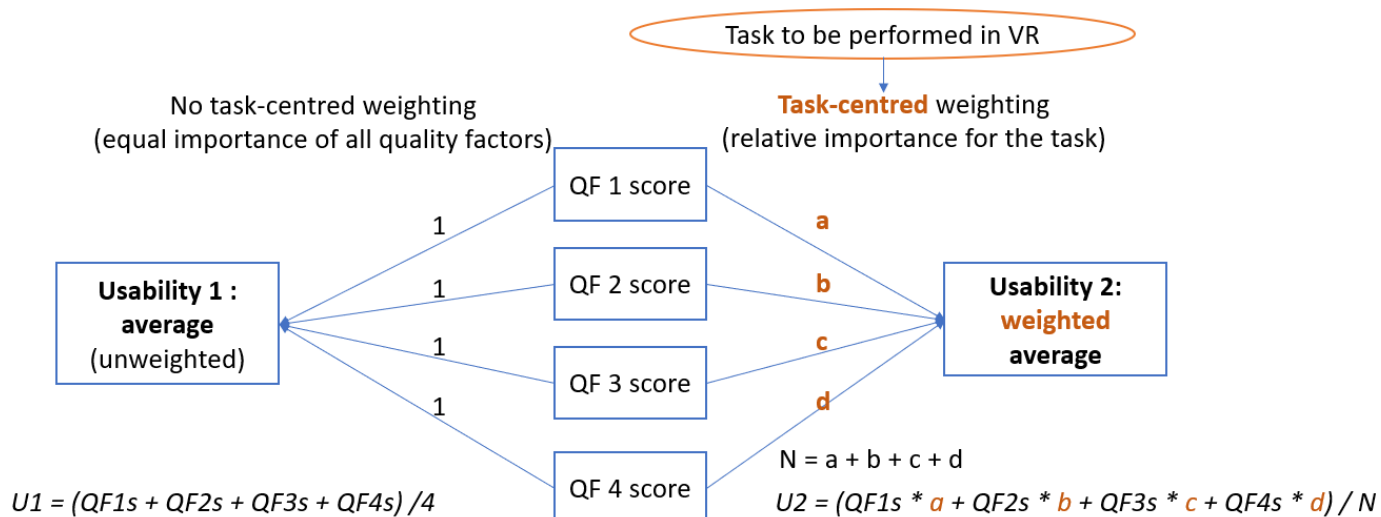

Figure 2. Quality factors and usability scores computation: average versus weighted average.

2.3. A Task-Centred Methodology for VR User Interactions: The Case Study of the Hazard Identification Task

2.3.1. Presentation of the Case Study and the Hazardous Situations

In this paper, we took as a case study the task of hazard identification, in a preventive context, i.e., before the beginning of the construction on a worksite. Indeed, thanks to the immersion that it provides, VR can improve the user task performance for this task that is usually performed by construction safety civil engineers on paper plans [34]. Nonetheless, the strength of the VR benefits may vary depending on the kinds of hazards. Indeed, hazards linked to the workers' behaviours benefit more from the VR immersion [49] than hazards related to worksite design issues such as trench collapse hazards, or to personal protection equipment issues, such as electrocution hazards. Consequently, in this study we focused on the two following kinds of hazards that directly involve workers: the falls ones and the struck-by ones—more specifically workers-vehicles collisions for these last ones. Figure 3 shows two examples of these kinds of hazards in our VR application.

Figure 3. Left: An worker in a fall hazard situation. **Right**: An worker in a struck-by hazard situation.

With these two kinds of hazards, our user task here is directly related to the workers' behaviours, and thus all the hazardous situations would involve a worker in our study. As a positive side-effect, this would help our participants to signal hazards in a unified way by always targeting a worker and not another element of the environment such as a vehicle. However, a potential negative side-effect might have been for this task to become a "search-a-worker" task: to avoid this, we put in our VR application some workers in both safe and hazardous situations. Similarly, to prevent false positives, we put basic personal protection equipment to all the virtual workers, so they could only be identified as in hazardous situations for the following reasons: the absence of external—not personal—protections such as guardrails or barriers, their proximity to a moving vehicle, or the use of inappropriate material on scaffolding such as pallets. Figure 4 shows on the top a worker in a safe situation thanks to guardrails, and on the bottom another one in a struck-by hazardous situation.

Figure 4. A worker in a safe situation on the top and another in a hazardous situation on the bottom.

2.3.2. Step 1: Hierarchical Analysis of the Hazard Identification Task

To design our VR user interactions for the hazard identification task, first we followed our hierarchical task analysis step that consists in decomposing this task into subtasks until reaching primitive tasks such as move, place, orient, etc. Through this analysis, we obtained the following primitive tasks:

- to move in the whole worksite;
- to activate the functionality to signal workers in hazardous situations;
- to target workers in hazardous situations;
- to place/orient the machines present on the worksite.

The primitive task "to move in the whole worksite"—i.e., exploration of the worksite—is one of the key subtasks for a successful hazard review since a bad locomotion would imply potential misses in terms of hazard detection, for example by going too fast. This is why we propose in this study to **focus on this exploration of the worksite subtask**, and thus on the **design** of its related user **navigation interaction**.

2.3.3. Step 2: Determination of Proposals of VR Interaction Techniques

For each primitive task, we used our rule-based system to determine proposals of VR interaction techniques, according to the characteristics of these primitive tasks. We obtained the following results:

- for the subtask "activate the functionality to signal workers in hazardous situations": physical/interface button interaction technique;
- for the subtask "target workers in hazardous situations": virtual raycast pointer or virtual hand go-go interaction technique;
- for the subtask "place/orient the machines present on the worksite": world-in-miniature interaction technique;
- for the subtask "explore the worksite": **two** potential **navigation interaction techniques** with one unfixed value (selection target mode).

Indeed, **for the worksite exploration subtask**, we obtained two proposals of navigation interaction techniques, which share the same following characteristics values: the control by the user of the direction in which to move without any restrictions, the user control of the acceleration through fits and starts actions, and a navigation speed in the human speed order of magnitude, since the users need to collect information for identifying hazard during the navigation. These two proposals differ from only one characteristic, the selection mode of the target for the next navigation point, each proposal having one of the two potential different values for this characteristic: the direct selection mode value—on the 3D space, and the indirect one—e.g., on a 2D map or through a list of zones. Based on that, we proposed the two following concrete implementations for our VR application prototypes:

- a free pointing steering technique with a **direct selection mode** on the 3D space, with smooth translation at a human speed to the targeted place;

- a free pointing steering technique with an **indirect selection mode** on a 2D map, with smooth translation at a human speed to the targeted place.

These two VR navigation interaction techniques consist for the user in targeting through a virtual raycast a place to move towards, and then in smoothly travelling to this place. The direct navigation technique allows a user to select the place or target to move towards by pointing on the ground of the 3D space of the worksite. The indirect navigation technique allows a user to select the place or target to move towards by pointing on a 2D map that represents the worksite. Figure 5 shows on the left a user moving with the direct technique, and on the right another user moving with the indirect technique.
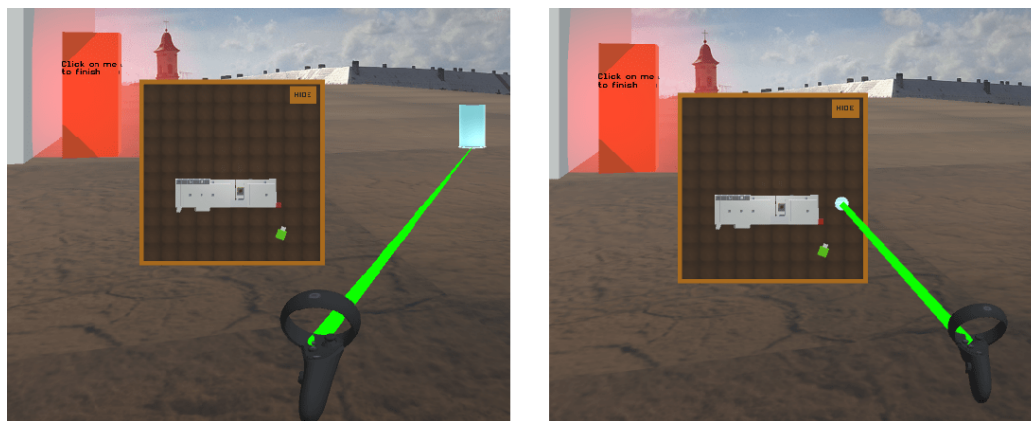
Figure 5. Left: "Direct target" navigation on the ground. **Right**: "Indirect target" navigation on a map.

2.3.4. Step 3: Creation of VR Prototypes and Task-Centred Evaluation Approach

For each of the navigation interaction techniques obtained from our step 2, we needed to develop a prototype of VR application. It should be highlighted that in our two prototypes, the interaction techniques related to all the other subtasks than the worksite exploration were also implemented thanks to step 2 results, such as hazardous situation targeting and tagging subtasks. We developed these VR prototypes in the Unity3D 2019 game engine, and the building models that we used here had been previously exported from Autodesk Revit 2018 BIM software.

To choose the VR devices to be used in this study, our main constraint was their accessibility to the experiment participants. Commonly, this accessibility issue in VR mainly comes from cost restrictions on the VR equipment, depending on where the experiment is conducted: available laboratory equipment, outdoor experiment restrictions, etc. In our case, we had to face two main restriction aspects, due to the full lockdown of the laboratories in 2020, because of the COVID-19 pandemic: our devices and data collection had to support a full remote usage, and our participants had to own their personal device. Thus, we chose to develop our VR prototypes for two different kinds of devices: computer desktops and VR head-mounted displays (HMD). We thought having two device possibilities should increase our participants' number. Therefore, we built four VR prototypes in total, one for each of the two interaction techniques for our two kinds of devices.

Our first available option for our participants was to use a desktop application on a computer. To allow for its remote distribution, we distributed our application online through a webpage, accompanied with all the required instructions and questionnaires. Our application was displayed on web browsers using the WebGL technology, and had been optimised for computer desktop screens. We proposed this modality to address our accessibility issues with this easily-accessible option. Additionally, pointing interaction techniques could be easily reproduced in this kind of 3D non-immersive environment, allowing us to get a correct evaluation of our navigation techniques.

The second option that we provided to our users was an HMD application. This option was viable thanks to an initiative of IEEE VR community members that had taken VR equipment outside of their laboratories during the lockdown and had proposed to share

remote experiments [67]. We also distributed our HMD application through a webpage with all the instructions and questionnaires. Participants had to download and install our application on their HMD to use it, allowing thus for having subjects that would perform their task in a 3D immersive environment despite the lockdowns. To maximise our participants number, we built several versions of our VR application for different HMDs: the HMD Oculus Quest, the Oculus Rift and the HTC Vive. This was possible since all these HMDs share the same key characteristics that we needed for this study: the tracking of the user head that allows for free head movements during the hazard identification review, and the presence of tracked hand controllers that our users could use for our navigation and selection interaction techniques.

Figure 6 shows a user navigating in the desktop application thanks to a mouse pointer on the left, and a user navigating in the HMD application using a HMD controller on the right.
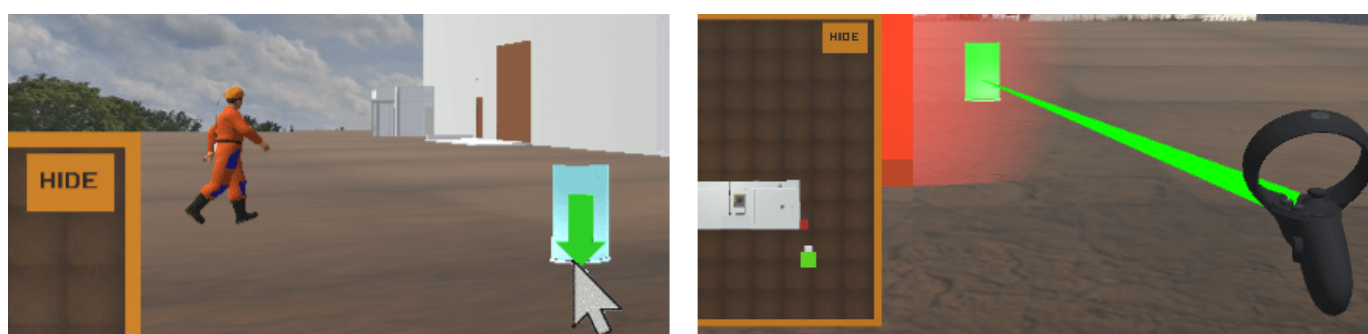
Figure 6. **Left**: A user using the desktop application. **Right**: A user using the HMD application.

Then, for our usability evaluation with users, different quality factors can be taken into account in the case of a navigation interaction. Bowman et al. [61] proposed the following list of quality factors for navigation interactions:

- the rapidity/speed during a travel;
- the precision/accuracy, i.e., the proximity of the arrival point compared to the desired target;
- the spatial awareness, i.e., "the user's implicit knowledge of disposition and orientation within the environment during and after travel";
- the ease of use of the interaction technique for a novice;
- the ease of learning of the interaction technique for a novice;
- the information gathering, i.e., "the user's ability to actively obtain information from the environment during travel";
- the presence felt by the users in the virtual environment thanks to the navigation interaction.

Nonetheless, Bowman et al. also noticed that depending on the user task and its context, some quality factors would be more relevant than others to evaluate the usability of an interaction technique [61]. As suggested in our methodology, a preliminary selection of the relevant quality factors for our specific case study and navigation interaction can be done through an interview with some AEC professionals. First, the professionals noted the importance of the information gathering quality factor because of the context of hazard identification, and considered the spatial awareness and ease-of-use quality factors as complementary factors. Indeed, our navigation interaction technique should allow for a great inspection of the virtual worksite when moving—information gathering factor, should not be complex so the users can focus on their inspection and not on navigating— easy to use factor, and should not make the users disoriented in order to let them locating spatially the hazards—spatial awareness factor. Moreover, they noted that, with large distances to be travelled during the worksite inspection and without any specific need of precision, the accuracy quality factor could be discarded, whereas the rapidity quality

factor could be kept. They nonetheless remarked the lower importance of the rapidity quality factor since the quality and the completeness of the hazard identification was more important than its rapidity. Finally, it was hard to consider in our study the presence quality factor since we were also using some non-immersive devices.

To sum up, in the case of our navigation interaction and exploration task, the following quality factors selected for our user study were:

- rapidity;
- information gathering ability;
- ease-of-use;
- spatial awareness.

Usability scores would be computed for each quality factor from the measures taken during the user study. Following our methodology, weights that represent the quality factor importance would be used for such computation, and in this study, they would be attributed by our participants through several weighting techniques that we explain in our user study protocol.

2.4. A User Study about Our Task-Centred Evaluation Approach on the Hazard Identification Case Study and Its Navigation Interaction in VR

2.4.1. Subjects

For our user study, we recruited voluntary unpaid subjects. Our experiment was available through the Internet without specific material requisites, except for the HMD participants. Oppositely, in terms of knowledge and skills requisites, performing the hazard identification task should have required for our participants to be safety engineers or construction workers. However, to avoid such limitation, we provided on the webpage of the experiment an informative document about the kinds of hazards that must be identified here. Similarly, to avoid computer skills prerequisites, a complete tutorial about the commands for navigating and tagging the workers in danger was provided. This allowed users with any kind of background to participate in our experiment.

Thanks to this design for our study, 34 different subjects participated to the experiment on the desktop version of our application, and 28 succeeded in performing correctly their task—the 6 remaining did not entirely finish the experiment. Data collected through a pre-questionnaire confirmed the presence of a large variety of profiles in terms of previous VR experience and AEC and safety knowledge. These subjects had mainly been recruited from the personal and professional networks of the authors. In the case of the HMD version of the application, half of the participants were VR experts from the IEEE VR community [67], and the other half were non-experts in VR. In total, 12 subjects participated to the HMD experiment with either the Oculus Quest, Oculus Rift or HTC Vive HMD.

2.4.2. Protocol and Design of the Experiment

Our experiment followed a within-subject design to evaluate the usability of our two navigation interaction techniques during the global task of hazard identification. Our participants had thus to perform this task twice, once with each navigation technique. For that, we built two scenarios with fall and struck-by hazardous situations in a unique virtual worksite. In these scenarios, some workers were positioned into different places of the worksite. Some of them were in safe placed, whereas the other ones were either close to a moving machine without any elements that could define a restricted walking area, or close to a high difference of height levels without any guardrails.

After having accepted with informed consent, the participants had to follow the following protocol, which was entirely explained on our experiment webpages. First, general explanations were given about the purpose of the experiment. Then, participants had to read an informative document about the hazards that would exist on the virtual worksite. After that, a tutorial about the VR application commands was given: navigation techniques, selection and manipulation techniques to target a worker, etc. Next, participants had to fill a short pre-experiment questionnaire about their previous experience and

knowledge about the VR technology, the hazard identification and mitigation and the AEC industry processes. Next to this form, they also had to attribute weights to the quality factors used in this usability evaluation. It was explained to the participants what a quality factor was, and short definitions of all the quality factors were given. They had to evaluate which quality factors were more important than the others ones by distributing weighting points.

Then, each participant had to perform a training in the application using the tagging worker interaction techniques and the navigation ones. Finally, each participant was able to pursue the hazard identification task in the application, either through the desktop or the HMD one, once with each of our two navigation interaction techniques. The order between the navigation interaction techniques and the scenarios—i.e., the different sets of hazardous situations were randomised and counterbalanced.

2.4.3. Measures for Each Quality Factor and Scores Computation

For each of our four quality factors—rapidity, information gathering, spatial awareness, and ease-of-use, measures were taken during the experiment. Based on our remote design for this experiment, with participants in full autonomy, first priority had been given to non-invasive and automatic measures. Inspired by previous experiments about navigation technique usability [58,61], we proposed to take the following measures, shown in Table 1.

Table 1. Quality factors and measures for navigation interaction during hazard identification task.

Quality Factor	Associated Measures
Rapidity	Navigation time (in seconds)
Information gathering	Number of hazards marked + time spent to identify (in seconds)
Situation awareness	Time spent for camera orientation (except when moving, in seconds)
Ease-of-use	Number of clicks required for navigating in all the worksite

From the data collected for each of these measures, we computed for each participant their personal scores for the four quality factors. These scores came from the normalisation of the data obtained from all the participants, after having processed the outlying values. We indeed computed these scores by applying a linear transformation of the data, with the value 1 attributed to the "best" value—e.g., the lower time, the higher number of hazards detected, etc., and 0 to the "worst" one. In the case of the information gathering quality factor—the only quality factor with two measures here, we computed an intermediate score for each of our two measures, and then we took the average of these two scores. Finally, this normalisation from 0 to 1 for all the scores allowed us to have the same scale for all of them, before computing the final usability scores, also on a 0 to 1 scale.

2.4.4. Weight Attribution to Quality Factors

In this study, to test our hypothesis we needed to compute both unweighted and weighted usability scores. To get these weights in this experiment, we asked the participants to attribute them through a pre-experiment questionnaire before using our VR application. Moreover, we used here two different techniques to obtain these relative importance weights, either by comparing all the quality factors importance at the same time, or by pair-comparison. By doing so, we expected that our users would be comfortable with at least one of these two techniques, and would succeed in attributing weights according to the importance they perceived for each quality factor in the context of this hazard identification task.

In the first technique, called "simple" weighting, participants had to define the weighting by distributing weights on four sliders from 0 to 6, as shown in Figure 7. Each value was directly reflecting the importance given to each quality factor. In the second technique, called the "cross-weighting" technique [68,69], they had to define the weights using a cross table, comparing the quality factors by pairs. Here, for each pair of quality factors,

participants had to select which quality factor was, as for them, the most important between both, and then they had to attribute a relative weight of 0, 1, or 2. A weight of 2 meant that the selected quality factor was "much more important" than the other, 1 that it was "moderately more important", and 0 meant that the two quality factors had the same importance. Figure 8 shows the interface provided in our webpage for the cross-weighting technique. Finally, pairwise weights must be summed to obtain a total weight for each quality factor and thus to have a weighted usability formula similar to the one obtained with the simple-weighting technique.
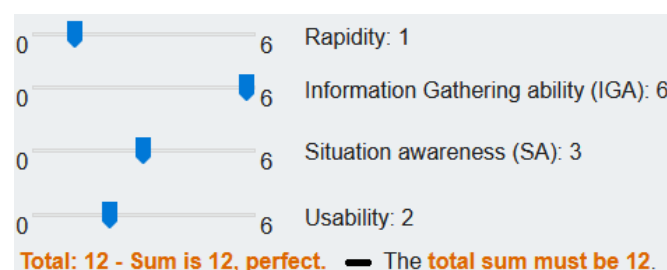

Figure 7. Simple weighting technique with sliders—global quality factors comparison.

Quality factors	Rapidity	Information Gathering ability (IGA)	Situation awareness (SA)	Usability
Rapidity	x	IGA ∨ 2 ∨	SA ∨ 1 ∨	Rapidity ∨ 1 ∨
Information Gathering ability (IGA)	x	x	IGA ∨ 2 ∨	IGA ∨ 1 ∨
Situation awareness (SA)	x	x	x	SA ∨ 0 ∨
Usability	x	x	x	x

Figure 8. Cross weighting technique with a cross-table—quality factors pair comparisons.

In total, we used in this user study four different ways to compute our global usability scores: (1) the unweighted way to get the "unweighted usability scores" without following our task-centred approach, (2) the user "simple-weighting" way, (3) the user "cross-weighting" way, and (4) an expert weighting way to get "weighted usability scores" following our task-centred evaluation approach. This expert weighting had been attributed before the experiment thanks to the following comments made by the AEC professionals about the different quality factors: the main importance of information gathering, the medium importance of the situation awareness and of the ease of use, and the lack of importance for the rapidity. This resulted in the following expert weighting: 0 for the rapidity, 4 for the information gathering, for the situation awareness 2, and for the ease of use 2—computed here using the cross-weighting technique.

3. Results

3.1. Prior Processing on the Datasets Obtained with Desktop and HMD Applications

3.1.1. Desktop Application: Results Dataset and Significant Effect of Order on Scores

We first needed to check on our 56-entries desktop application dataset the potential effects of the order—the number of the personal session when a user performed (1 or 2), or of the scenario—the set of hazards to be identified—factors on our quality factor and usability scores. Indeed, as our main objective was to study the interaction technique factor effect, we aimed to avoid having any significant effect of other factors. To do this analysis, an ANOVA could be run on our data with the interaction technique, order, and scenario factors: we started thus by verifying the required ANOVA assumptions. First, we applied the Shapiro-Wilk's test on our series of data—the four quality factor scores and the four global usability scores. This test returned p-values higher than a level of significance of $\alpha = 0.01$ or of $\alpha = 0.05$ depending on the series, which ensured that our data followed a normal distribution for all our scores. Then, we applied the Levene's test: it returned

p-values higher than a level of significance of $\alpha = 0.01$ or of $\alpha = 0.05$ depending on the series, which showed the homogeneity of the variance between the different possibilities of interaction and scenario combination for all the users. Finally, our three observed factors were independent by their nature—the navigation interaction techniques, the session order and the scenarios.

Therefore, we ran an ANOVA for the three aforementioned factors on all our scores, and Table 2 shows the p-values obtained for the scenario and order factors. Interaction effects between the different factors, which are not displayed here, were not significant on a level of significance of $\alpha = 0.05$. First, it appears that the scenario factor has not had a significant effect on any scores. However, we can observe a significant effect of the order factor on several scores results with p-values inferior to a level of significance of $\alpha = 0.05$. This is because participants improved significantly their scores in their second session, no matters the interaction technique used.

Table 2. Desktop application: repeated-measures ANOVA on quality factors and usability scores with interaction technique, scenario, and order factors.

Scores	Order Factor p-Value	Scenario Factor p-Value
Rapidity	0.959	0.369
Information gathering	0.329	0.380
Situation awareness	0.001 **	0.495
Ease of use	0.0003 ***	0.893
Unweighted usability	0.003 **	0.959
Simple-w. usability	0.0006 ***	0.949
Cross-w. usability	0.004 **	0.829
"Expert" usability	0.00004 ***	0.408

Note: *** indicates that p is <0.001, ** indicates that p is <0.01, and * indicates that p is <0.05 for all tables.

Thereupon, we ran again an ANOVA on the scores obtained by the participants during the first use of our application, with the interaction technique and scenario factors, to check if there were no residual significant effect of the scenario factor with the order fixed to the first session. Table 3 reveals no significant effects of scenario on our data. As a result, for a cleaner analysis of the interaction technique effect, we decided to use then this 28-entries subset of data with only the users' first session results. Thus, next figures and discussions about the desktop application results would be based on this "between-subjects dataset".

Table 3. Desktop application: between subjects ANOVA on quality factors and usability scores with interaction technique and scenario factors.

Scores	Scenario Factor p-Value
Rapidity	0.970
Information gathering	0.393
Situation awareness	0.500
Ease of use	0.765
Unweighted usability	0.520
Simple-w. usability	0.552
Cross-w. usability	0.868
"Expert" usability	0.276

3.1.2. HMD Application: Results Dataset

For our HMD application dataset, we also needed to check the potential effects of the order and the scenario factors on our scores, by running an ANOVA with these two factors and the interaction technique one. Before that, we checked if our data respected the required ANOVA assumptions. First, our data followed a normal distribution for each of our data series, according to the Shapiro-Wilk's test that returned for all the series a p-value

higher than a level of significance of $\alpha = 0.05$. About the sphericity of the data, we checked it graphically by observing similar variances for each group. Indeed, the Mauchly's test can only be run when there are more than two levels for the factors, and with only two levels—our case here—the literature said that sphericity necessarily holds for effects. Finally, the three observed factors were independent by their nature.

Then, after running this ANOVA with these three factors on our scores, no significant effect of the scenario of hazardous situations or of the session order factors have been detected, as shown in Table 4. As a result, the results shown and analysed in the next figures and discussions would take into account the scores present in the complete within-subject HMD application dataset.

Table 4. HMD: repeated-measures ANOVA on quality factors and usability scores with interaction technique, scenario, and order factors.

Scores	Order Factor *p*-Value	Scenario Factor *p*-Value
Rapidity	0.644	0.389
Information gathering	0.244	0.581
Situation awareness	0.076	0.010
Ease of use	0.229	0.538
Unweighted usability	0.249	0.213
Simple-w. usability	0.164	0.174
Cross-w. usability	0.132	0.445
"Expert" usability	0.055	0.265

3.2. Desktop Application Results

First, Figure 9 show the users' scores results for each individual quality factor, from left to right: the rapidity, the information gathering, the situation awareness and the ease of use. Next, Figure 10 shows the usability scores results that we computed with the four different ways that we described, from left to right: the scores without using weights, then the ones using the weights given by the users with the simple-weighting technique, next the ones that rely on the cross-weighting technique used by the users, and finally the ones with the expert weighting.

Doing a brief visual analysis, it can be noticed in Figure 9 that it seems that the participants performed better in terms of rapidity with the direct navigation interaction technique than with the indirect one. Moreover, it seems to be the same trend with the information gathering and situation awareness scores. Finally, about the ease of use, it seems to be the opposite trend, more moderately, with better scores with the indirect technique. About the global usability scores, it seems that Figure 10 shows that the usability was better with the direct mode for the unweighted scores, and the simple weighted scores. For the cross-weighted and expert-weighting scores, this trend does not appear so clearly; the usability might be similar with the two interaction techniques.
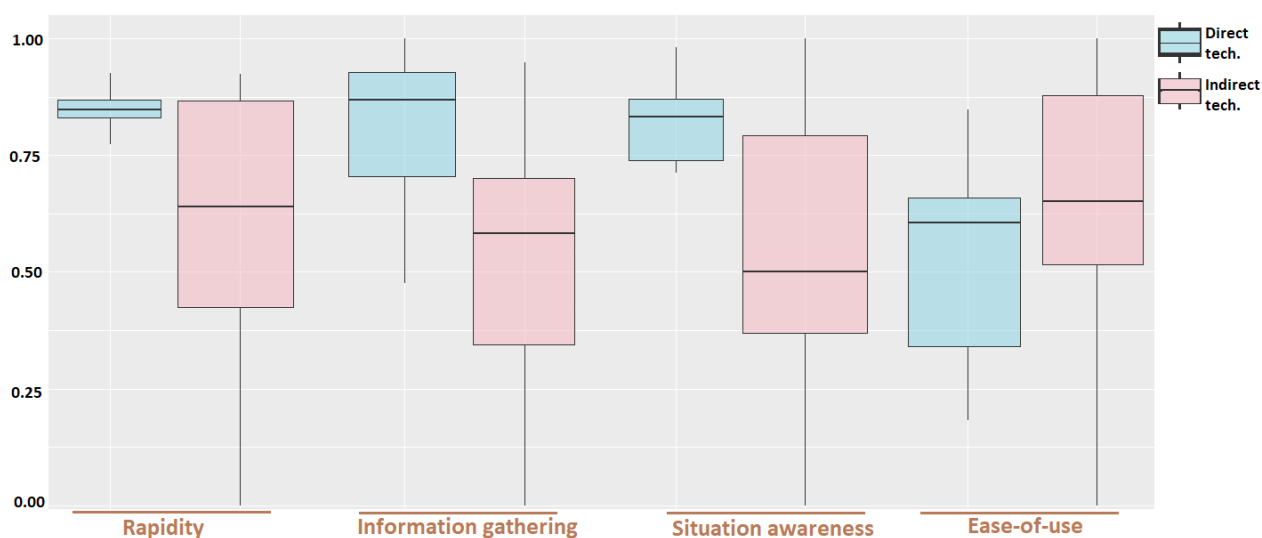
Figure 9. Quality factors scores distribution—between desktop subjects data.
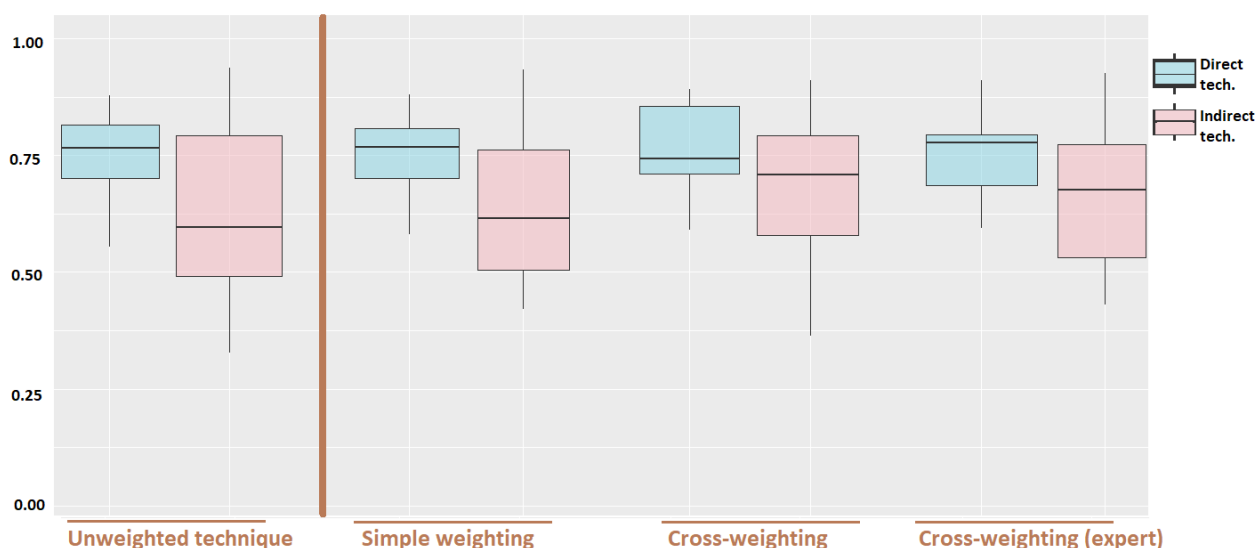

Figure 10. Usability scores results distribution—between desktop subjects data.

3.3. HMD Application Results

First, Figure 11 shows the distribution of the results obtained for, from left to right, the rapidity, information gathering, situation awareness, and ease of use quality factor scores. Next, the results for our four computation of usability scores can be observed in Figure 12, from left to right: the scores without using weights, then the ones using the weights given by the users with the simple-weighting and the cross-weighting techniques, and finally the ones with the "expert" weighting.

Doing a brief visual analysis, in Figure 11 it seems that the participants performed better in terms of rapidity with the direct navigation interaction technique. Then, it seems to be the same trend with the information gathering scores and the situation awareness scores. Regarding the ease of use, it tends to be the opposite trend with better scores with the indirect technique. Finally, about the usability scores, Figure 12 may show that the usability was better with the direct technique, for the unweighted scores on the left. However, for the other (weighted) usability scores present in the figure, this trend tends to not be as pronounced as with the unweighted ones.
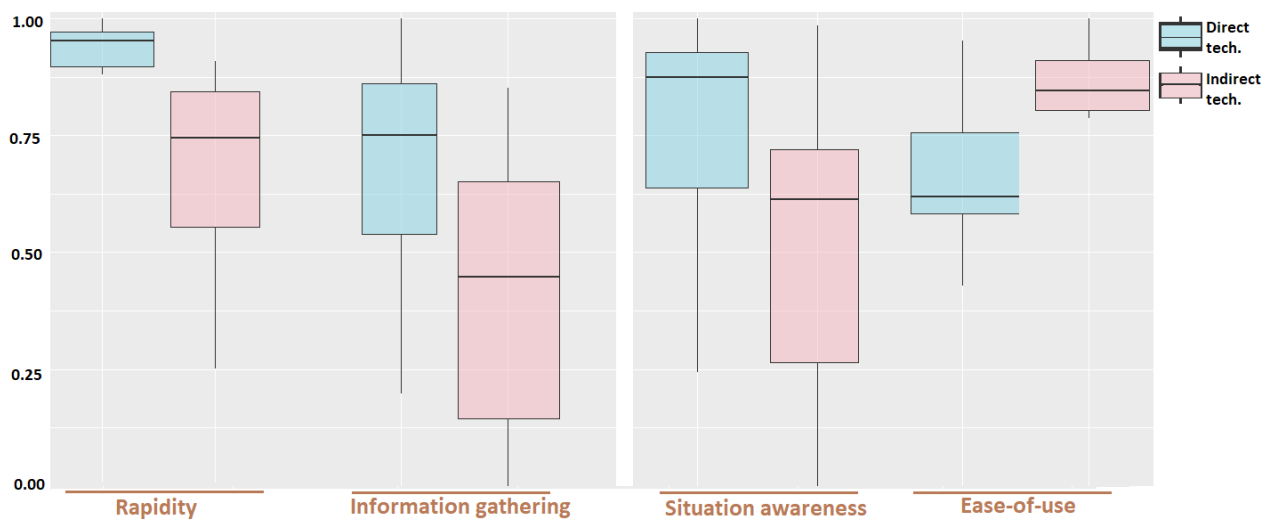
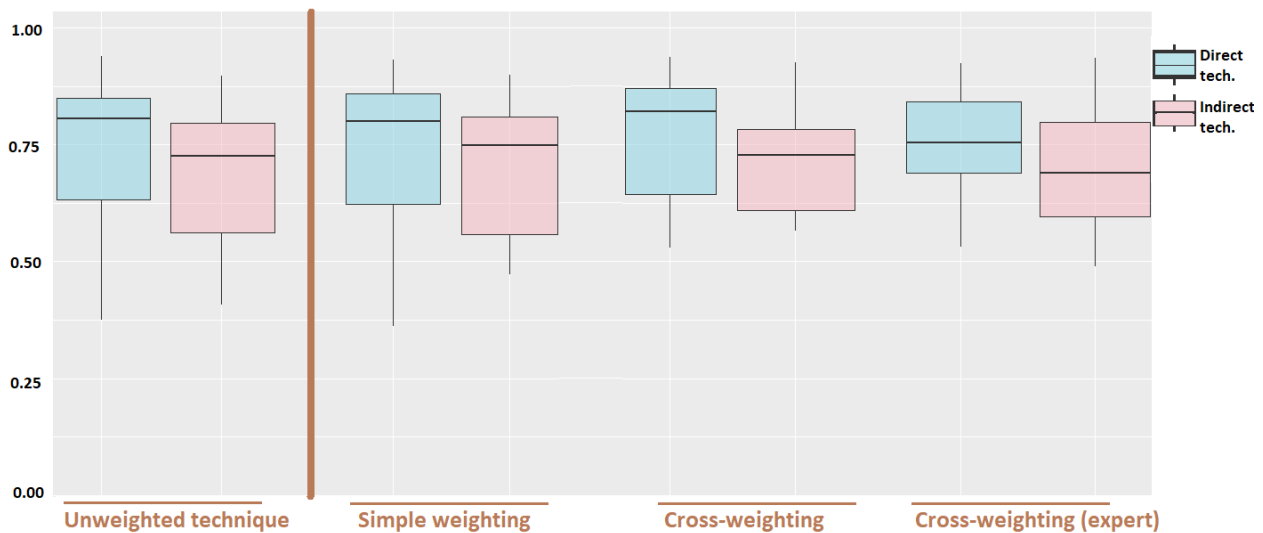**Figure 11.** Quality factors scores distribution—within HMD subjects data.

Figure 12. Usability scores distribution—within HMD subjects data.

4. Discussion

4.1. Statistical Analysis on Desktop Application Usability Scores and Discussion

For our desktop application dataset, Table 5 (on the top) shows the p-values obtained from the two-way ANOVA conducted on all our quality factor scores, considering the interaction technique and scenario factors, and Table 6 (on the bottom) shows the ones obtained for all our usability scores. Scenario p-values revealed no significant effect as shown previously in Table 3. Focusing then on the interaction technique factor, Table 5 (on the top) shows its related p-values for the rapidity, information gathering, situation awareness and ease of use scores, and Table 6 (on the bottom) for the four usability scores.

Table 5. Desktop application: between subjects—two-way ANOVA on quality factors scores with interaction technique and scenario factors.

Scores	int. tech. Factor p-Value
Rapidity	0.006 **
Information gathering	0.002 **
Situation awareness	0.013 *
Ease of use	0.170

Table 6. Desktop application: between subjects—two-way ANOVA on usability scores with interaction technique and scenario factors.

Scores	int. tech. Factor *p*-Value
Unweighted usability	0.048 *
Simple-w. usability	0.032 *
Cross-w. usability	0.090
"Expert" usability	0.085

This statistical analysis confirmed the results observed previously in the figures. First, the interaction technique factor does have a significant effect on the rapidity, information gathering and situation awareness scores with *p*-values of 0.006, 0.002, and 0.013, respectively, on a level of significance of $\alpha = 0.05$. By looking at the results, the best scores for these quality factors are with the direct interaction technique. Moreover, there is no significant effect of the interaction technique on the "ease of use" scores, even if a subtle trend can be noticed with better scores with the indirect technique.

Then, this analysis shows that the interaction technique factor does have a significant effect on the unweighted usability score with a *p*-value of 0.048 on a level of significance of $\alpha = 0.05$. With the simple-weighting, this is the same, with a *p*-value of 0.032 on a level of significance of $\alpha = 0.05$. However, with the cross-weighting usability scores, both with the weighting defined by the users and the expert one, there is no significant effect of the interaction technique with *p*-values of 0.090 and 0.085, respectively, on a level of significance of $\alpha = 0.05$. **These last results tend to confirm our hypothesis about the effect of our task-centred approach on the interaction technique usability results**. Indeed, the interpretation of the usability results for our two navigation interaction techniques with weighted scores is **different and opposed** to the one done for the unweighted scores results.

In total, with this desktop application, two over three weighted computations led to an interpretation that is different compared to the one with the unweighted computation, as expected from our hypothesis. Nonetheless, due to this unexpected result with the computation from the simple-weighting technique, some future research should be done on the techniques used to define the quality factors weights. In this experiment, such differences could be explained by the fact that, even if it may be more complex for the users, the cross-weighting technique allowed for a weighting distribution by the users that was closer to the expert one than with the simple-weighting. Indeed, in our experiment, it was expected for the participants to attach the greatest importance to the information gathering quality factor, since this is the most important quality factor for this task of exploration and identification, by definition. However, recomputing the weights with a total of 12 points in all the cases for both weighting techniques, on average the participants gave more importance to this quality factor with the cross-weighting technique—average: 5.20—than with the simple-weighting—average: 4.16. Similarly, it was expected to obtain the lowest possible weights for the least important quality factor for this task, the rapidity; once again, on average our users gave less importance to it with the cross-weighting technique—average: 1.15—than with the simple-weighting technique—average: 1.67.

4.2. Statistical Analysis on HMD Application Usability Scores and Discussion

For our HMD application dataset, Table 7 (on the top) shows the *p*-values obtained from our three-way ANOVA on all our quality factor scores, considering the interaction technique, order and scenario factors, and Table 8 (on the bottom) shows the ones obtained for all our usability scores. Order and scenario *p*-values revealed no significant effect as shown previously in Table 4. Table 7 (on the top) shows the interaction technique *p*-values for the rapidity, information gathering, situation awareness and ease of use scores, and Table 8 (on the bottom) for the four usability scores.

Table 7. HMD application: repeated-measures ANOVA on quality factors scores with interaction technique, scenario, and order factors.

Scores	int. tech. Factor *p*-Value
Rapidity	0.043 *
Information gathering	0.0002 ***
Situation awareness	0.0006 ***
Ease of use	0.0009 ***

Table 8. HMD application: repeated-measures ANOVA on usability scores with interaction technique, scenario, and order factors.

Scores	int. tech. Factor *p*-Value
Unweighted usability	0.048 *
Simple-w. usability	0.200
Cross-w. usability	0.638
"Expert" usability	0.114

This statistical analysis confirmed the results observed graphically on the previous figures. First, the interaction technique factor does have a significant effect on the rapidity, information gathering, situation awareness and ease of use scores with *p*-values of 0.043, 0.0002, 0.0006, and 0.0009, respectively, on a level of significance of $\alpha = 0.05$. Looking at the scores, the best ones are with the direct interaction technique for the rapidity, information gathering, and situation awareness quality factors, whereas for the ease of use the best scores are with the indirect interaction technique.

Then, the analysis shows that the interaction technique factor does have a significant effect on the unweighted usability score with a *p*-value of 0.048 on a level of significance of $\alpha = 0.05$. However, with all our three different weighted scores proposed here—the simple-weighting, the cross-weighting by the users, and the expert weighting—there is not a significant effect of the interaction technique factor on the usability scores, with *p*-values of 0.200, 0.638, and 0.114, respectively, on a level of significance of $\alpha = 0.05$. This means that with the HMD application, in our user study three over three weighted techniques gave opposite results and interpretations for the evaluated interaction techniques, compared to the unweighted evaluation. **These results confirm our hypothesis about the effect of our task-centred approach on the interaction technique usability results and their interpretation, and its difference with a non-task-centred approach**.

5. Conclusions and Future Work

Virtual reality simulations are paramount in the workflow of numerical engineering, before physical actions in reality such as on a worksite in the case of the AEC industry. In this context, we presented in this paper our task-centred methodology to design and evaluate virtual reality user interactions. To improve this numerical workflow, we created this methodology to guide and facilitate the design of VR applications by VR end users. Indeed, our methodology provides them with the benefit of a formalised approach for the design of VR user interactions, which they can follow without having VR expertise. Through a guided analysis of a user task and its characteristics, a professional in the field to which this task belongs can obtain proposals of VR user interaction designs with our semi-automated system in very few steps. About the evaluation of these VR user interaction designs, our task-centred approach allows for the formalisation of the users' needs, which improves as a result the quality of the VR user interaction designs in terms of usability. Moreover, this formalisation should also help the VR end users to reduce the number of required iterations when creating a VR application for their specific purposes.

After that, we showed in this paper how we applied our task-centred methodology on the case study of the hazard identification task: construction of a model of the task, de-

termination of proposals for the design of VR user interaction, and user evaluation. For the subtask of worksite exploration, we obtained two possible designs of navigation interaction techniques. These two interaction techniques could then be prototyped, evaluated and compared in terms of usability in our evaluation step, which mainly relies on applying a task-related weight on each quality factor used in the current evaluation. In this study, we proposed to compare the usability results and interpretations of our task-centred evaluation with weights, with the ones without applying such different weights and letting all the quality factors at the same level of importance. Our hypothesis was that our task-centred evaluation based on weights would give usability results that could be interpreted in a significantly different way in comparison with a non-task-centred evaluation without any task-related weights.

To run this study and its related user evaluation, we built VR applications with our two different navigation interaction techniques. Due to the sanitary crisis context, we conducted this experiment remotely and our users participated through remote modalities—either a desktop application or an HMD application. We took different measures for each quality factor about this navigation interaction, and then we computed both weighted usability scores—with three different formulas—and unweighted ones. Considering both our desktop and HMD modalities, five over six weighted usability scores computed here give an interpretation of the results that is opposed to the one obtained using unweighted usability scores. Indeed, our unweighted scores show a significantly different usability for our two navigation interaction techniques, whereas our weighted scores reveal no significant differences in terms of usability. These results verified thus our hypothesis about the difference between task-centred and non-task-centred evaluations of VR user interactions. This validates the importance of following an approach that takes into account the expertise of the professionals in the field to which a task belongs to create a VR application for performing this task, and, therefore, our task-centred approach is validated.

As future work, it would be interesting first to test and evaluate in a user study the different ways that we used to obtain the quality factors weights, since one way led to one weighted computation with different results than our other five ones. Indeed, a dedicated experiment could be conducted on the evaluation of the weighting procedures themselves—simple and cross weightings—with, for example, some criteria of usability, of understanding how to attribute weights, etc. This would be an important verification in order to know which weighting technique should preferably be used in a task-centred evaluation. Then, our results about our hypothesis should be confirmed by running similar experiments on other case studies with different tasks, interaction techniques and quality factors. Additionally, thanks to these other experiments, a machine learning model could be trained with the aim of attributing automatically the quality factors weights according to the user task. Finally, for on-site applications, which are based on the augmented reality technology instead of the virtual reality technology, our task-centred methodology could be adapted for augmented reality interactions, and studies could be conducted applying this modified methodology.

Author Contributions: Conceptualization, P.R.; Data curation, P.R.; Methodology, P.R.; Software, P.R.; Supervision, R.L., F.D., P.F., J.T.H. and F.M.; Validation, P.R. and R.L.; Visualization, P.R.; Writing—original draft, P.R.; Writing—review & editing, P.R., R.L., F.D., P.F., J.T.H. and F.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available at https://github.com/pierreraimbaud/webp/tree/master/data (accessed on 30 April 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AEC	Architecture Engineering Construction
BIM	Building Information Modelling
VR	Virtual Reality
AR	Augmented Reality
HMD	Head-Mounted Display

References

1. Xu, X.; Ma, L.; Ding, L. A Framework for BIM-Enabled Life-Cycle Information Management of Construction Project. *Int. J. Adv. Robot. Syst.* **2014**, *11*, 126. [CrossRef]
2. Gambatese, J.A.; Hallowell, M. Factors that influence the development and diffusion of technical innovations in the construction industry. *Constr. Manag. Econ.* **2011**, *29*, 507–517. [CrossRef]
3. Nowotarski, P.; Paslawski, J. Industry 4.0 Concept Introduction into Construction SMEs. *IOP Conf. Ser. Mater. Sci. Eng.* **2017**, *245*, 052043. [CrossRef]
4. Sacks, R.; Eastman, C.; Lee, G.; Teicholz, P. *BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers*; John Wiley & Sons: Hoboken, NJ, USA, 2018.
5. Eastman, C.M. Preliminary Report on a System for General Space Planning. *Commun. ACM* **1972**, *15*, 76–87. [CrossRef]
6. Chen, H.; Hou, L.; Zhang, G.K.; Moon, S. Development of BIM, IoT and AR/VR technologies for fire safety and upskilling. *Autom. Constr.* **2021**, *125*, 103631. [CrossRef]
7. Bullinger, H.J.; Bauer, W.; Wenzel, G.; Blach, R. Towards user centred design (UCD) in architecture based on immersive virtual environments. *Comput. Ind.* **2010**, *61*, 372—379. [CrossRef]
8. Behzadan, A.H.; Kamat, V.R. Interactive Augmented Reality Visualization for Improved Damage Prevention and Maintenance of Underground Infrastructure. In *Building a Sustainable Future*; American Society of Civil Engineers: Reston, VA, USA, 2009 ; pp. 1214–1222; doi:10.1061/41020(339)123. [CrossRef]
9. Kumar, A.; Sharma, S.; Goyal, N.; Singh, A.; Cheng, X.; Singh, P. Secure and energy-efficient smart building architecture with emerging technology IoT. *Comput. Commun.* **2021**, *176*, 207–217. [CrossRef]
10. Raimbaud, P.; Lou, R.; Merienne, F.; Danglade, F.; Figueroa, P.; Hernandez, J.T. BIM-based Mixed Reality Application for Supervision of Construction. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; pp. 1903–1907. [CrossRef]
11. Zaker, R.; Coloma, E. Virtual reality-integrated workflow in BIM-enabled projects collaboration and design review: A case study. *Vis. Eng.* **2018**, *6*, 4. [CrossRef]
12. Harirchian, E.; Aghakouchaki Hosseini, S.E.; Jadhav, K.; Kumari, V.; Rasulzade, S.; Işık, E.; Wasif, M.; Lahmer, T. A review on application of soft computing techniques for the rapid visual safety evaluation and damage classification of existing buildings. *J. Build. Eng.* **2021**, *43*, 102536. [CrossRef]
13. Harirchian, E.; Jadhav, K.; Kumari, V.; Lahmer, T. ML-EHSAPP: A prototype for machine learning-based earthquake hazard safety assessment of structures by using a smartphone app. *Eur. J. Environ. Civ. Eng.* **2021**, 1–21. [CrossRef]
14. Jin, R.; Zou, P.X.; Piroozfar, P.; Wood, H.; Yang, Y.; Yan, L.; Han, Y. A science mapping approach based review of construction safety research. *Saf. Sci.* **2019**, *113*, 285–297. [CrossRef]
15. Alves Dias, L. Construction Safety in Portugal. In *Cib Report*; A. A. Balkema Publishers: Brookfield, VT, USA, 1996; pp. 43–56.
16. Duff, A. Management and operative safety improvement: A goal for the whole organisation. In Proceedings of the International Conference of CIB Working Commission W99, Lisbon, Portugal, 4–7 September 1996.
17. Kamardeen, I. 8D BIM modelling tool for accident prevention through design. In Proceedings of the 26th Annual ARCOM Conference, Association of Researchers in Construction Management Leeds, Leeds, UK, 6–8 September 2010; Volume 1, pp. 281–289.
18. Zhou, Z.; Goh, Y.M.; Li, Q. Overview and analysis of safety management studies in the construction industry. *Saf. Sci.* **2015**, *72*, 337–350. [CrossRef]
19. Zhang, S.; Sulankivi, K.; Kiviniemi, M.; Romo, I.; Eastman, C.M.; Teizer, J. BIM-based fall hazard identification and prevention in construction safety planning. *Saf. Sci.* **2015**, *72*, 31–45. [CrossRef]
20. Zhang, S.; Teizer, J.; Lee, J.K.; Eastman, C.M.; Venugopal, M. Building information modeling (BIM) and safety: Automatic safety checking of construction models and schedules. *Autom. Constr.* **2013**, *29*, 183–195. [CrossRef]
21. Albert, A.; Pandit, B.; Patil, Y. Focus on the fatal-four: Implications for construction hazard recognition. *Saf. Sci.* **2020**, *128*, 104774. [CrossRef]

22. Kim, K.; Kim, H.; Kim, H. Image-based construction hazard avoidance system using augmented reality in wearable device. *Autom. Constr.* **2017**, *83*, 390–403. [CrossRef]
23. Cheung, W.F.; Lin, T.H.; Lin, Y.C. A Real-Time Construction Safety Monitoring System for Hazardous Gas Integrating Wireless Sensor Network and Building Information Modeling Technologies. *Sensors* **2018**, *18*, 436. [CrossRef]
24. Yilmaz, F. Analysis of occupational accidents in construction sector in Turkey. *J. Multidiscip. Eng. Sci. Technol. JMEST* **2014**, *1*, 421–428.
25. Martínez-Aires, M.D.; López-Alonso, M.; Martínez-Rojas, M. Building information modeling and safety management: A systematic review. *Saf. Sci.* **2018**, *101*, 11–18. [CrossRef]
26. Suraji, A. Analysis of labour accident in the construction industry: The Indonesia experience. *J. TEKNIKA* **1997**, *8*, 16–20
27. Xiaer, X.; Dib, H.; Yuan, J.; Tang, Y.; Li, Q. Design for Safety (DFS) and Building Information Modeling (BIM): A Review. In Proceedings of the International Conference on Construction and Real Estate Management 2016, Guangzhou, China, 10–12 November 2017; pp. 69–80. [CrossRef]
28. Qi, J.; Issa, R.; Hinze, J.; Olbina, S. Integration of Safety in Design through the Use of Building Information Modeling. *Comput. Civ. Eng.* **2011**, 698–705. [CrossRef]
29. Gambatese, J.; Hinze, J. Addressing construction worker safety in the design phase: Designing for construction worker safety. *Autom. Constr.* **1999**, *8*, 643–649. [CrossRef]
30. Qi, J.; Issa, R.R.A.; Olbina, S.; Hinze, J. Use of Building Information Modeling in Design to Prevent Construction Worker Falls. *J. Comput. Civ. Eng.* **2014**, *28*, A4014008. [CrossRef]
31. Shiau, Y.R.; Lo, F.Y.; Ko, P.C. Early intervention mechanism for preventing electrocution in construction engineering. *Ind. Health* **2021**, *59*, 4–17. [CrossRef] [PubMed]
32. Chi, C.F.; Lin, S.Z. Classification scheme and prevention measures for caught-in-between occupational fatalities. *Appl. Ergon.* **2018**, *68*, 338–348. [CrossRef]
33. Hinze, J.W.; Teizer, J. Visibility-related fatalities related to construction equipment. *Saf. Sci.* **2011**, *49*, 709–718. [CrossRef]
34. Cooke, B.; Williams, P. *Construction Planning, Programming and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
35. Sulankivi, K.; Kähkönen, K.; Mäkelä, T.; Kiviniemi, M. 4D-BIM for construction safety planning. In Proceedings of the W099-Special Track 18th CIB World Building Congress, Salford, UK, 10–13 May 2010; pp. 117–128.
36. Cruz, O.G.D. Virtual Reality (VR): A Review on its Application in Construction Safety. *Turk. J. Comput. Math. Educ. TURCOMAT* **2021**, *12*, 3379–3393.
37. Fuchs, P.; Moreau, G.; Guitton, P. *Virtual Reality: Concepts and Technologies*; CRC Press: Boca Raton, FL, USA, 2011.
38. Slater, M.; Usoh, M.; Steed, A. Depth of Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* **1994**, *3*, 130–144. [CrossRef]
39. Mütterlein, J. The Three Pillars of Virtual Reality? Investigating the Roles of Immersion, Presence, and Interactivity. In Proceedings of the 51st Hawaii International Conference on System Sciences, Waikoloa Village, HI, USA, 2–6 January 2018; pp. 1–10. [CrossRef]
40. Li, X.; Yi, W.; Chi, H.L.; Wang, X.; Chan, A.P. A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Autom. Constr.* **2018**, *86*, 150–162. [CrossRef]
41. Kazar, G.; Comu, S. Effectiveness of Serious Games for Safety Training: A Mixed Method Study. *J. Constr. Eng. Manag.* **2021**, *147*, 04021091. [CrossRef]
42. Sacks, R.; Perlman, A.; Barak, R. Construction safety training using immersive virtual reality. *Constr. Manag. Econ.* **2013**, *31*, 1005–1017. [CrossRef]
43. Zhao, D.; Lucas, J. Virtual reality simulation for construction safety promotion. *Int. J. Injury Control Saf. Promot.* **2015**, *22*, 57–67. [CrossRef] [PubMed]
44. Zhao, D.; Lucas, J.; Thabet, W. Using virtual environments to support electrical safety awareness in construction. In Proceedings of the 2009 Winter Simulation Conference (WSC), Austin, TX, USA, 13–16 December 2009; pp. 2679–2690. [CrossRef]
45. Fang, Y.; Teizer, J.; Marks, E. A Framework for Developing an As-built Virtual Environment to Advance Training of Crane Operators. In *Construction Research Congress 2014*; American Society of Civil Engineers: Reston, VA, USA, 2014; pp. 31–40. [CrossRef]
46. Xu, Z.; Zheng, N. Incorporating Virtual Reality Technology in Safety Training Solution for Construction Site of Urban Cities. *Sustainability* **2021**, *13*, 243. [CrossRef]
47. Joshi, S.; Hamilton, M.; Warren, R.; Faucett, D.; Tian, W.; Wang, Y.; Ma, J. Implementing Virtual Reality technology for safety training in the precast/prestressed concrete industry. *Appl. Ergon.* **2021**, *90*, 103286. [CrossRef] [PubMed]
48. Sacks, R.; Whyte, J.; Swissa, D.; Raviv, G.; Zhou, W.; Shapira, A. Safety by design: Dialogues between designers and builders using virtual reality. *Constr. Manag. Econ.* **2015**, *33*, 55–72. [CrossRef]
49. Perlman, A.; Sacks, R.; Barak, R. Hazard recognition and risk perception in construction. *Saf. Sci.* **2014**, *64*, 22–31. [CrossRef]
50. Hadikusumo, B.; Rowlinson, S. Integration of virtually real construction model and design-for-safety-process database. *Autom. Constr.* **2002**, *11*, 501–509. [CrossRef]
51. Lucena, A.F.E.; Saffaro, F.A. Guidelines for exploring construction sites in virtual reality environments for hazard identification. *Int. J. Occup. Saf. Ergon.* **2020**, 1–10. [CrossRef]
52. Jerald, J. *The VR Book: Human-Centered Design for Virtual Reality*; Morgan & Claypool: San Rafael, CA, USA, 2015.

53. Wann, J.; Mon-Williams, M. What does virtual reality NEED? Human factors issues in the design of three-dimensional computer environments. *Int. J. Hum. Comput. Stud.* **1996**, *44*, 829–847. [CrossRef]

54. Sutcliffe, A.G.; Poullis, C.; Gregoriades, A.; Katsouri, I.; Tzanavari, A.; Herakleous, K. Reflecting on the design process for virtual reality applications. *Int. J. Hum. Interact.* **2019**, *35*, 168–179. [CrossRef]

55. Longo, L. Experienced mental workload, perception of usability, their interaction and impact on task performance. *PLoS ONE* **2018**, *13*, e0199661. [CrossRef]

56. Nielsen, J.; Levy, J. Measuring Usability: Preference vs. Performance. *Commun. ACM* **1994**, *37*, 66–75. [CrossRef]

57. Frokjaer, E.; Hertzum, M.; Hornbaek, K. Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'00, Hague, The Netherlands, 1–6 April 2000; Association for Computing Machinery: New York, NY, USA, 2000; pp. 345–352. [CrossRef]

58. Bowman, D.A.; Hodges, L.F. Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *J. Vis. Lang. Comput.* **1999**, *10*, 37–53. [CrossRef]

59. Chen, J.; Bowman, D. Domain-Specific Design of 3D Interaction Techniques: An Approach for Designing Useful Virtual Environment Applications. *Presence* **2009**, *18*, 370–386. [CrossRef]

60. Gabbard, J.L.; Hix, D.; Swan, J.E. User-Centered Design and Evaluation of Virtual Environments. *IEEE Comput. Graph. Appl.* **1999**, *19*, 51–59. [CrossRef]

61. Bowman, D.A.; Koller, D.; Hodges, L.F. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In Proceedings of the IEEE 1997 Annual International Symposium on Virtual Reality, Albuquerque, NM, USA, 1–5 March 1997; pp. 45–52.

62. Bowman, D.A.; Gabbard, J.L.; Hix, D. A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence Teleoperators Virtual Environ.* **2002**, *11*, 404–424. [CrossRef]

63. Poupyrev, I.; Weghorst, S.; Billinghurst, M.; Ichikawa, T. A framework and testbed for studying manipulation techniques for immersive VR. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, Lausanne, Switzerland, 15–17 September 1997; pp. 21–28.

64. Annett, J. Hierarchical task analysis. *Handb. Cogn. Task Des.* **2003**, *2*, 17–35.

65. Arns, L.L. A New Taxonomy for Locomotion in Virtual Environments. Ph.D. Thesis, Iowa State University, Ames, IA, USA, 2002; doi:10.31274/rtd-180813-14259. [CrossRef]

66. De Feo, G.; De Gisi, S. Using an innovative criteria weighting tool for stakeholders involvement to rank MSW facility sites with the AHP. *Waste Manag.* **2010**, *30*, 2370–2382. [CrossRef] [PubMed]

67. Steed, A.; Ortega, F.R.; Williams, A.S.; Kruijff, E.; Stuerzlinger, W.; Batmaz, A.U.; Won, A.S.; Rosenberg, E.S.; Simeone, A.L.; Hayes, A. Evaluating Immersive Experiences during Covid-19 and Beyond. *Interactions* **2020**, *27*, 62–67. [CrossRef]

68. Saaty, T.L. What is the analytic hierarchy process? In *Mathematical Models for Decision Support*; Springer: Berlin/Heidelberg, Germany, 1988; pp. 109–121.

69. Saaty, T.L. Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* **2008**, *1*, 83–98. [CrossRef]