Article

The Moderating Effect of Debriefing on Learning Outcomes of IVR-Based Instruction: An Experimental Research

Qinna Feng, Heng Luo *, Wenhao Li *, Ying Chen and Jiakai Zhang

Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan 430079, China; fengqinna@mails.ccnu.edu.cn (Q.F.); chenying2020@mails.ccnu.edu.cn (Y.C.); jkzhang@mails.ccnu.edu.cn (J.Z.)
* Correspondence: luoheng@mail.ccnu.edu.cn (H.L.); whenhowlee@mail.ccnu.edu.cn (W.L.)

Featured Application: Debriefing is an essential component of IVR-based instruction with varying moderating effects on knowledge acquisition and behavioral performance.

Abstract: With its ability to afford immersive and interactive learning experiences, virtual reality has been widely used to support experiential learning, of which the learning effectiveness is promoted by the instructional component of debriefing. The current literature on debriefing mainly focuses on the traditional learning contexts while little is known on its effectiveness in immersive virtual reality (IVR) learning environments. Based on the theories of experiential learning and debriefing, this study designed a debriefing strategy based on simulated learning experience and investigated its effectiveness on knowledge and behavioral learning in an IVR learning program, using a randomized controlled trial with 77 elementary students from Hubei province in China. The study results support the efficacy of IVR on improving knowledge acquisition and behavioral performance, and reveal a significant moderating effect of debriefing on the effectiveness of IVR learning environments. The study confirms the critical role of debriefing in IVR-based instruction and provides theoretical and practical implications for the design and implementation of effective IVR learning environments.

Keywords: virtual reality; debriefing; instructional design; learning outcomes; experimental design

1. Introduction

Advancement in artificial intelligence technologies has shaped the landscape of education profoundly in recent years, with technological innovations such as virtual learning environments and adaptive learning systems drastically improving the quality of teaching and learning at all levels [1]. Among these, virtual reality (VR) merits our special attention due to its unique cognitive and psychological benefits [2,3]. VR is a digital environment based on computer simulation and 3D modeling technology that closely simulates the physical environment in terms of visual, audio, tactile, and interactive experience [4]. According to the degree of immersion, the application of VR in instruction can be divided into desktop-based virtual reality (DVR) and immersive virtual reality (IVR). The literature has shown that IVR can boost learning engagement, knowledge transfer, empathy, and learner agency [5,6]. In recent years, with the advancement in function and the reduction in costs, IVR technology has gradually been accepted in the field of education and applied to various disciplines such as sciences [7], humanities [8], and second language acquisition [5]. IVR can also simulate the authenticity and complexity of dangerous situations while ensuring learners' physical safety. It is, thus, also widely used in various safety education and training programs, such as fire escape [9], mine safety [10], and pedestrian safety [11].

Although IVR has more advanced technical features such as stereoscopic display, depth of field (DoF) modeling and tracking, high-resolution 3D rendering, and natural interaction, its actual effect in educational practice has been disappointing. The substantial increase in immersion and interactivity has not yielded superior learning performance. For example, Wu et al. (2020) conducted a meta-analysis of 35 experiments from 2013 to

2019 and found that the overall learning effect of IVR (Hedge's g) was 0.24, which was smaller than the overall effect of DVR ($g = 0.41$–0.51) [12]; Luo et al. (2021) reached a similar conclusion through meta-analysis of 149 experimental studies [2]. One possible reason for this phenomenon is that IVR isolates learners from physical reality completely, which causes a lack of interaction with teachers and peers during the learning process, and is, thus, not conducive to the deployment of teaching strategies. Compared with the technical affordances of IVR, the selection and application of proper teaching strategies play a more critical role in improving the effectiveness of IVR-based instruction [12].

Debriefing is a simple and easy-to-use teaching strategy that can potentially improve the learning effect of IVR. This strategy guides learners to recall, analyze, and evaluate their own or others' behavioral decisions in virtual scenarios through visualized and shared cognitive processes, thereby promoting reflective learning and knowledge construction in experiential learning [13]. Debriefing is widely practiced in military training and medical education, and has achieved positive learning results. For example, Moldjor et al. (2015) showed that listening to experienced military personnel's reflective reports before conducting military training and rescue operations could increase trust and understanding among team members [14]; Brown et al. (2018) revealed that one-on-one debriefing after simulation training obtained better results in subsequent medical knowledge tests [15]. In recent years, researchers have applied debriefing strategies in various educational contexts, including teacher training [16], transnational cultural education [17], and psychological therapy [18], and have reported significantly improved learning outcomes. However, most debriefing strategies described in the literature were implemented in traditional classrooms [19] with non-immersive simulation interventions [15,16], with adults such as college students and corporate workers as target learners [16]. There is a lack of rigorous empirical evidence concerning whether or not the debriefing strategy is also applicable to the IVR learning environment, whether it can improve the learning performance of younger learners, and to what extent.

In summary, informed by the experiential learning theory, IVR-based instruction has the potential to improve students' learning performance and experience. However, the current literature of VR-based instruction suffers from three limitations. First, many studies focused on the technical features of the IVR learning environment and ignored the implementation and application of teaching strategies outside IVR, such as debriefing. Second, although the effectiveness of debriefing strategies has been verified in the traditional classroom, it is still unclear whether the relevant conclusions are applicable to the IVR learning environment, as the moderating effect of the debriefing on IVR learning performance lacks rigorous and systematic investigation. Third, the target population of debriefing strategies were usually adult learners, children as potential IVR learners have not been sufficiently investigated. In view of those limitations, we designed and implemented a debriefing strategy for an IVR-based instruction as part of a child pedestrian safety education project and investigated the moderating effect of this teaching strategy on the IVR learning outcomes in terms of knowledge acquisition and behavioral performance.

2. Literature Review

2.1. Experiential Learning Supported by IVR

Experiential learning theory states that learning is a process in which learners connect education, tasks, and personal experience in a complex or dynamic learning space. This learning space supports four learning behaviors: concrete experience, reflective observation, abstract hypotheses, and active testing [20]. Experiential learning values the creation of learning environments and emphasizes situational cognition and knowledge construction based on collaboration and reflection. With the rapid development of VR technology and popularization of HMDs, IVR technology can engage learners in a virtual learning environment characterized by a high degree of immersion, interaction, and imagination, thereby supporting the development of experiential learning activities. Researchers have developed various software applications to study the VR-mediated learning experiences in

different fields, at different stages, and with various learning outcomes [8]. The results in general indicate that the experiential learning experiences afforded by IVR have a positive impact on students' learning outcomes, yet the effect sizes were less than satisfactory [2,12]. A possible reason is that many IVR instructional interventions ignore the role of the teacher to facilitate learning outside the IVR environment, which results in a lack of meaningful interaction and feedback between learner and teacher, thus causing learning gains to vary.

2.2. Debriefing

Debriefing is also known as guided reflection, which is an indispensable part of experiential learning. It is a teaching strategy widely used in simulation-based learning [21]. Debriefing refers to the process of an instructor guiding learners to recall and evaluate their learning experiences with prompt questions after a simulated experience. The goal of debriefing is to help learners to gain deeper understanding of their decisions in simulation and promote knowledge construction and transfer though critical reflection or discussion [22]. The debriefing strategy was first applied to military exercises during World War II [23] and was later adopted by the fields of psychology and education. Debriefing in the field of education is different from the recall and report practice in military simulations as it has a systematic process with essential instructional elements. It is usually combined with concrete experiential learning activities that serve as instructional contexts for social interaction and shared inquiry, thereby overcoming the limitations of didactic lecture methods [24].

Lederman (1992) has summarized the seven essential elements of debriefing in simulation-based experiential learning: the facilitator/debriefer, the participants, the experience, the impact of that experience, the recollection of the experience, the mechanisms for reporting on the experience, and the time needed to process the experience. These seven elements interact with each other to form the three stages of the reflection process: the introduction to systematic reflection and analysis, the intensification and personalization of the analysis of the experience, and the generalization and application of the experience [20]. From the perspective of the seven elements and three stages, debriefing in education not only implies self-recall or reporting on the experience, but also emphasizes the use of effective strategies to guide learners to carry out personalized reflection and transfer reflection results to practical applications.

Research on debriefing in the literature has mostly been seen in simulation-based training in traditional brick-and-mortar contexts, such as onsite medical training [15], military drill [14], and psychological therapy [18]. The results indicate that this strategy not only improves learning outcome but can also stimulate higher learning motivation, increased self-efficacy, and improve learning satisfaction. However, the literature lacks statistical evidence based on rigorous experimental research as to whether these research conclusions on debriefing are also applicable to virtual learning environments, especially experiential learning afforded by IVR [11].

2.3. Knowledge and Behavioral Learning

While knowledge and behavior are considered independent constructs that differ in their cognitive and educational origins [25], they are also intricately interrelated as learning objectives for IVR-based instruction: A wide range of knowledge and skills must be obtained for students to make correct behavioral decisions to accomplish the learning tasks in IVR [26], and the diagnosis of behavioral performance can also inform the selection of content and methods for knowledge training [11].

Anderson divides knowledge into declarative and procedural knowledge according to the different operations for processing that knowledge and the way it is stored in the human brain [27]. The former is closely related to the cognitive functions of recall, recognition, and recitation, and the latter is related to the cognitive functions of comprehension, analysis, and application. Procedural knowledge not only requires acquisition and comprehension of knowledge but also needs to transform that knowledge into behavioral changes and

accomplish cross-context knowledge transfer and theory application to solve authentic problems [28].

Our review of the literature has identified two potential issues regarding the evaluation of knowledge acquisition and behaviors performance in IVR. The first issue is measurement validity: The use of surveys and other self-reporting methods is prevalent to evaluate learning effects, which lacks accuracy and objectivity. For example, Fung et al. (2019) used questionnaires to understand students' feedback towards IVR learning [29]. The second issue is the lack of systematic comparison of knowledge and behavioral learning outcomes in IVR. For example, Taranilla et al. (2019) used test questions to measure the degree to which knowledge was mastered by students [8], and Schwebel et al. (2018) tested the effect of IVR on improving children's pedestrian performance by coding their street-crossing behaviors. Although the research results showed that learners improved their level of knowledge or behavioral performance, there was a lack of comparison in terms of the degree of improvement between the two [30].

2.4. Research Questions

In summary, debriefing is a key instructional component of experiential learning that promotes meaningful reflection and interaction. However, the effect of debriefing in the IVR learning environment lacks systematic investigation, especially concerning its actual effect size on the learning outcomes. In addition, existing research lacks comparative analysis of the two types of learning outcomes of IVR-based instruction: knowledge acquisition and behavioral performance. Consequently, based on an IVR intervention for pedestrian safety education, this study investigated the moderating effect of a designed debriefing strategy on children's knowledge and behavioral learning outcomes. The application of IVR in safety education can ensure the physical safety of children while allowing teachers to make a full diagnosis of children's risky pedestrian behaviors, which provides a foundation for the evaluation and measurement of pedestrian performance. Second, IVR can provide learners with concrete and immersive learning experiences, laying the foundation for subsequent individualized reflection activities. Specifically, the following questions guided our investigation:

1. Does the use of the debriefing strategy in the IVR learning environment have a moderating effect on knowledge acquisition and behavioral improvement?
2. Is there a difference between the influence of debriefing on students' knowledge acquisition and behavioral improvement in the IVR learning environment?

3. Methods

This study utilized a randomized experimental design to investigate the moderating effect of debriefing on IVR learning outcomes. The presence of the debriefing activity is the grouping variable that defined the treatment and control group. The dependent variables are the two types of IVR learning outcomes: knowledge acquisition and behavioral performance. The specifics of the research methods are elaborated on below.

3.1. Participants

To ensure sufficient power for the experiment design, this study calculated the sample size required for a repeated measurement variance mixed experiment to obtain a large effect ($\eta_p^2 > 0.14$). The results indicated that the sample size of this experiment (N = 77) is much larger than the minimum required sample size (N = 24). This experiment selected all of the students in the second and third grades of a school who agreed to participate in the experiment. These students came from towns and suburbs and had diverse family backgrounds and academic abilities. Before the start of the experiment, we had already obtained written informed consent from all the parents of the students. In addition, all of the students reported that their vision was normal or normal after putting on glasses, and stated that they had little or no experience with IVR. Students were randomly divided into two groups to participate in the experiment: the treatment group (n = 39) and the control

group (n = 38). There were 24 male students in the treatment group and 21 male students in the control group. Those students were also randomly assigned to five different IVR learning sites. Each site was operated by four research assistants who were responsible for informing students about how to walk in the virtual environment (Coordinator), ensuring the safety of the students during the IVR experience (Safety Guider), videotaping the whole process (Video Manager), and guiding students in the reflection (Debriefer). To minimize the potential individual influence of facilitation, all research assistants received week-long training on standardized facilitation protocols.

3.2. IVR Learning Environment

We used Unity 3D game engine and Oculus Rift development kit to create the IVR learning environment investigated in this study. This environment was a virtual road traffic scene constructed by simulating the road traffic in the city center. The virtual "avatar" and the vehicles behaved as in the real environment so that the learner could be deeply immersed in the virtual environment. For a detailed introduction to the IVR program, please refer to the research team's prior paper [11].

In the IVR learning environment, learners wore HMDs and held sensor controllers to allow free exploration within the tracking space. As shown in Figure 1, participants needed to complete three challenges at the three intersections designed in the IVR program: (1) When an avatar friend on the opposite side beckons the participant to join him, the participant needs to choose whether to cross the street immediately or wait for the next green light to cross at the pedestrian crossing (C1); (2) participants need to correctly interpret the meaning of a flashing green pedestrian light and make decisions regarding whether to cross a wide street (C2); (3) when the pedestrian light is green but a backing school bus is ready to drive away, the participant needs to choose whether to continue crossing the street or to wait for the bus to leave (C3). To successfully accomplish the three pedestrian challenges, participants need to demonstrate five correct pedestrian behaviors: B1—do not dash into the street; B2—cross the street when the pedestrian light is green; B3–when the pedestrian light is blinking green, do not cross the street when at the sidewalk or quickly cross it when in the middle of the crosswalk; B4—exam the incoming traffic when crossing the street; B5—be alert to the backing vehicle and give it priority over traffic light for safety.

3.3. Debriefing Strategy

This research designed an individual debriefing strategy based on an IVR learning environment. When the students completed the three crossroad challenges in the IVR environment, the instructor observed and recorded the participant's reactions and decisions they made in a timely manner. Then, according to the participant's decisions made, the instructor would give the students one-on-one individual guidance after the task was finished. Based on the three stages of debriefing proposed by Lederman (1992), this study proposed an individualized debriefing strategy that consisted of four stages: VR experience review, guided self-assessment, individualized feedback, and knowledge transfer [24].

As shown in Figure 2, at the first stage, the instructor asked students what they heard, saw, and felt during the IVR learning process, this helped students to systematically recall the experiences of going through the three pedestrian challenges. At the second stage, the instructor uncovered students' conception of traffic rules based on their performance in the IVR environment, and helped them to carry out self-examination and reflection. At the third stage, the instructor provided individualized feedback regarding the correct traffic knowledge relevant to students' risky behaviors. At the fourth stage, the instructor proposed one or two hypothetical situations and guided students to apply the learned traffic knowledge to solve novel pedestrian problems within the situations.
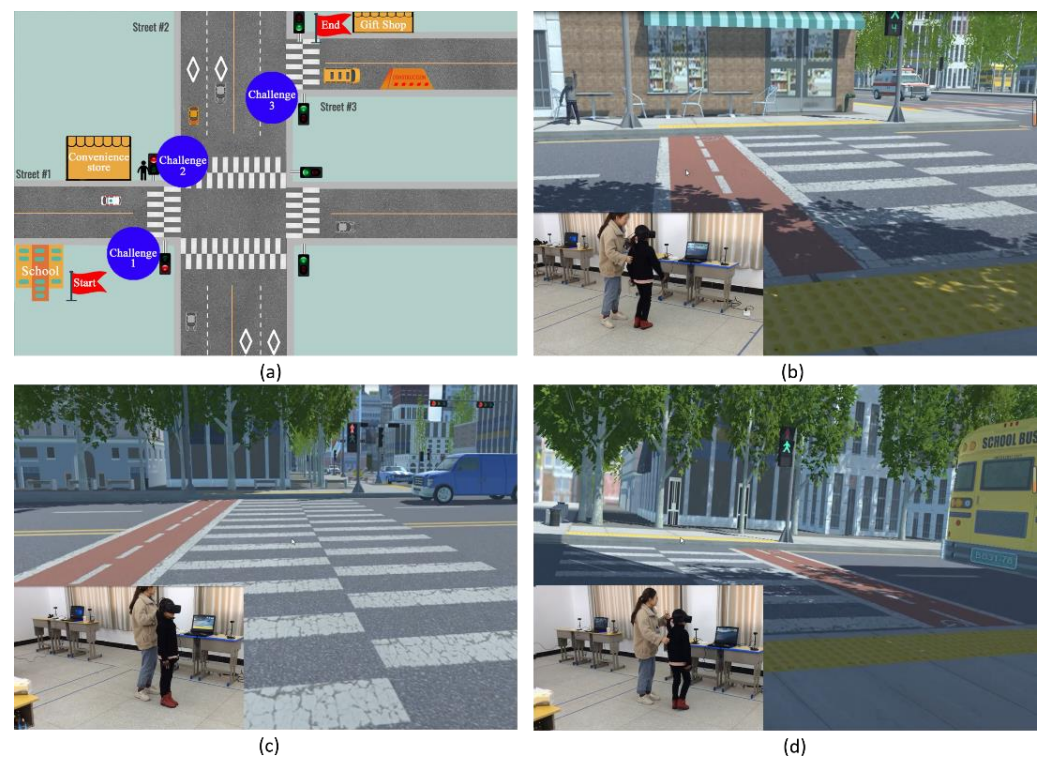
Figure 1. (**a**) visualization of three challenges in the VR program; (**b**) C1: participants need to cross the street safely to meet with an avatar friend; (**c**) C2: participants need to correctly interpret the meaning of a flashing green pedestrian light and make decisions regarding whether to cross a wide street; (**d**) C3: participants need to delay their street-crossing activity and wait for a backing school bus to drive away.
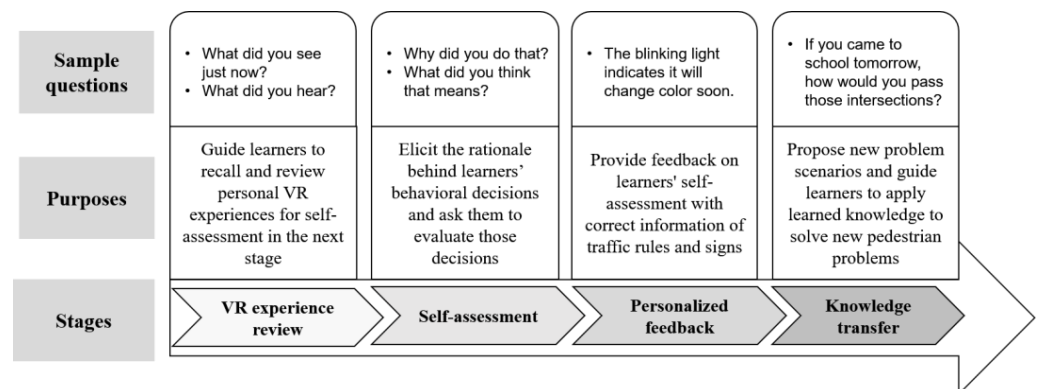

Figure 2. The instructional design of the debriefing strategy based on IVR experience.

3.4. Experiment Procedure

The study employed a two-factor mixed experimental design to examine the impact of IVR on children's learning performance and the moderating effect of the proposed debriefing strategy. The internal factor tested was time (before and after the IVR-based instruction), and the within-subject factor tested was the intervention (debriefing and no debriefing strategy). The experimental process consisted of four phases and six activities as shown in Figure 3, and lasted about 30 min.
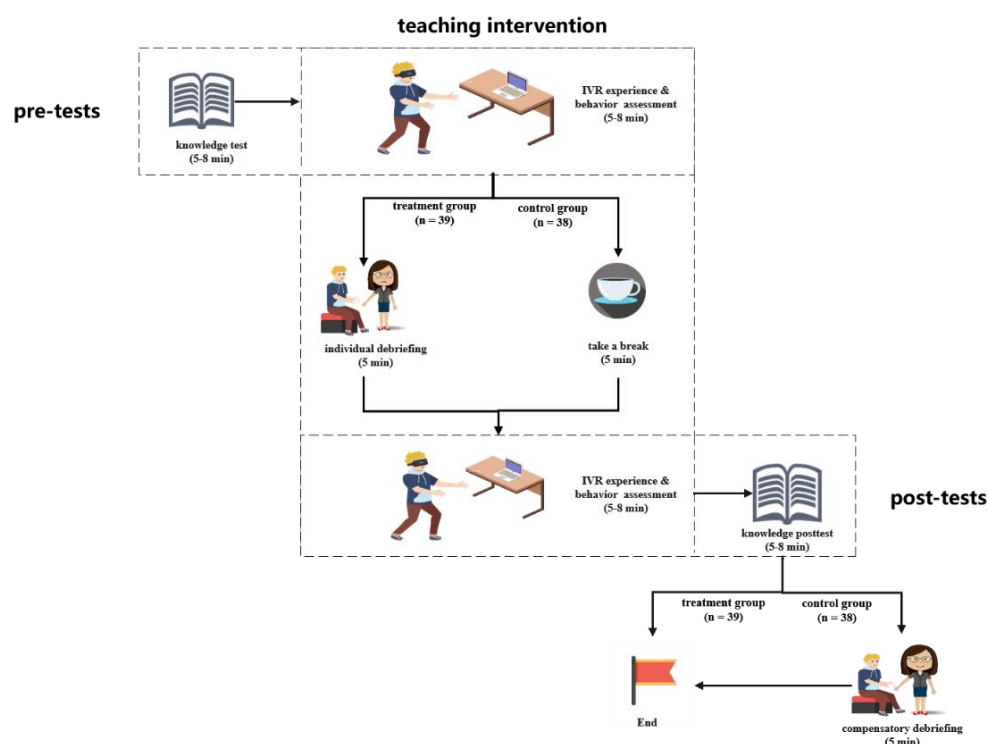
Figure 3. The general procedure of the experimental design.

The first phase included a knowledge pretest and a behavior pre-assessment. First, all participants were required to complete a paper-based knowledge test; afterwards, they had their first IVR learning experience by completing the three pedestrian challenges. The IVR learning experience also served as a behavior pre-assessment as all participants' pedestrian behaviors were video recorded for future analysis. Proceeding into the second phase, the debriefing session received by the treatment group was considered as an instructional intervention, which lasted for about 5 min. The control group did not participate in the debriefing session and took a 5-min break during this period.

The third phase was identical with the first phase. All participants had a second IVR learning experience, which also served as the behavior post-assessment, and then proceeded directly into the knowledge posttest that was the same as the knowledge pretest. The fourth phase was the interview. The researcher asked the participants about the IVR learning experience and what they had learned based on the first three phases, using questions to help them make a summary promptly and boost knowledge transfer, such as "How did your learning go today?" or "What will you do next time when you cross the street?" To ensure all participants obtained the same quality of pedestrian safety education, the control group also participated in the identical debriefing session after the experiment.

The entire experiment process was captured with video cameras, including participants' interactions with the IVR program and their performance in the debriefing sessions (if applicable). Additionally, we also used the computer screen recording function to record the first-person view of participants' behaviors within the virtual world created by the IVR.

3.5. Data Collection and Analysis

There were two main types of data collected in this study: the pedestrian knowledge test scores and pedestrian behaviors within the IVR program. The two types of data were used to measure participants' knowledge acquisition and behavioral performance. An independent sample *t*-test and factorial repeated measure ANOVA were conducted separately to examine the differences in the posttest scores and the moderating effect of debriefing. IBM SPSS 21 software was the selected tool for statistical analysis.

The instrument for collecting participants' knowledge acquisition was a paper-based knowledge test developed by the researchers (The knowledge test is accessible at https://www.doi.org/10.17632/tffr34c5fr.1, accessed on 2 November 2021). There were 10 single-choice questions in the knowledge test divided into two dimensions: interpretation (4 questions) and decision making (6 questions). Interpretation questions tested students' mastery of traffic rules, such as "What does zebra crossing/crosswalk mean?" and "When a pedestrian traffic light is green and flashing, what does it mean?". The decision-making questions were designed to measure students' ability to make correct pedestrian choices in the face of complex traffic situations, such as "Your soccer ball is rolling across the street, and there is a crosswalk 100 m in front of you. How should you get your soccer ball?" The full mark for the knowledge test was 10, with each item counting for one point.

The participants' pedestrian behavioral data were captured by both the onsite video recording and the first-person point of view screen capture. The video-recorded behavioral performances were analyzed and rated by two researchers using an observation rubric (see Appendix A) that dictated the following behavioral scoring rules: For a particular behavior, every time a participant made the decision on that behavior, the correct and incorrect decisions were marked as 1 and 0, respectively. Accordingly, the performance score for that behavior was measured by the percentage of correct decisions. For example, if a participant checked the traffic before crossing streets once out of three street-crossing instances, her final performance on this behavior (checking traffic) was calculated as 0.33 (1/3). Further, a participant's overall pedestrian performance was scored by dividing the total points she gained by the total number of behaviors observed. Consequently, the scores for the overall pedestrian performance as well as individual pedestrian behaviors were between 0 (0% correctness, failed completely) to 1 (100% correctness, flawless performance).

4. Results

4.1. Initial Analysis

An independent sample *t*-test was used to analyze whether there were differences in prior knowledge and personal characteristics between the treatment group and the control group. The results showed that both the knowledge and behavior scores met the homogeneity of variance, and there was no statistically significant group difference in the knowledge pretest scores ($t = 1.978$, $p = 0.052 \geq 0.05$) or the behavior pre-assessment scores ($t = 0.59$, $p = 0.629 \geq 0.05$). However, it is worth noticing that the initial gap in the knowledge test scores (MD = 1.12) was larger than that in behavioral performance (MD = 0.023). Additionally, the chi-square test results showed that there were no significant differences between the two groups in terms of gender ($\chi^2 = 0.6879$, $p \geq 0.05$) or grade ($\chi^2 = 0.013$, $p \geq 0.05$), so there was no significant difference in the prior knowledge level, grade, or gender distribution between the treatment group and the control group.

4.2. Moderating Effect of Debriefing

To investigate the effect of the IVR learning experience on students' knowledge acquisition and behavioral performance, a repeated measures ANOVA was performed to determine knowledge acquisition and behavior improvement of both groups before and after the IVR-based instruction, with key results listed in Table 1. Compared with the pretest scores, the posttest scores of both pedestrian knowledge and behavior improved significantly (MD $_{(knowledge)}$ = 1.41, MD $_{(behavior)}$ = 0.12, $p = 0.000 < 0.001$) for all participants. The general effect sizes (η_p^2) were 0.309 and 0.434, respectively, indicating that engagement in IVR experience could effectively assist students to acquire traffic knowledge and improve pedestrian behavior, regardless of the debriefing session. However, an interaction effect of debriefing and IVR experience was found on the increasement of behavior scores (F (1, 75) = 7.627, $p = 0.007$, $\eta_p^2 = 0.092$) rather than the knowledge test scores (F (1, 75) = 2.068, $p = 0.155$, $\eta_p^2 = 0.028$). The results indicated that the effectiveness of IVR on promoting behavioral performance was largely dependent on the implementation of the debriefing

strategy, but the impact of debriefing on knowledge acquisition was limited, suggesting a moderating effect.

Table 1. Within-subjects effects of the trials and the interaction with debriefing (N = 77).

Variable	Trials [a]	Within-Subjects Effects			Trial × Debriefing		
		F	p	$\eta_p{}^2$	F	p	$\eta_p{}^2$
Knowledge test scores	pretest: 5.840 (2.510) posttest: 7.220 (2.377)	32.640	0.000	0.309 **	2.068	0.155	0.028
Behavior assessment scores	pretest: 0.412 (0.171) posttest: 0.536 (0.211)	57.612	0.000	0.434 **	7.627	0.007	0.092 **

[a] The values depict means (standard deviations). ** The effect is significant at $p < 0.01$.

The moderating effect of debriefing on IVR learning is perhaps more clearly shown in Figure 4 when we further compared the effect sizes ($\eta_p{}^2$) of learning gains in knowledge and behavior between treatment and control groups. The effect sizes for knowledge acquisition and behavioral improvement were 0.368 and 0.611 for the treatment group, which were higher than those of the control group (0.240 and 0.222). This result indicates that the moderating effect of debriefing on behavior performance was far greater than that on knowledge acquisition in IVR-based instruction.
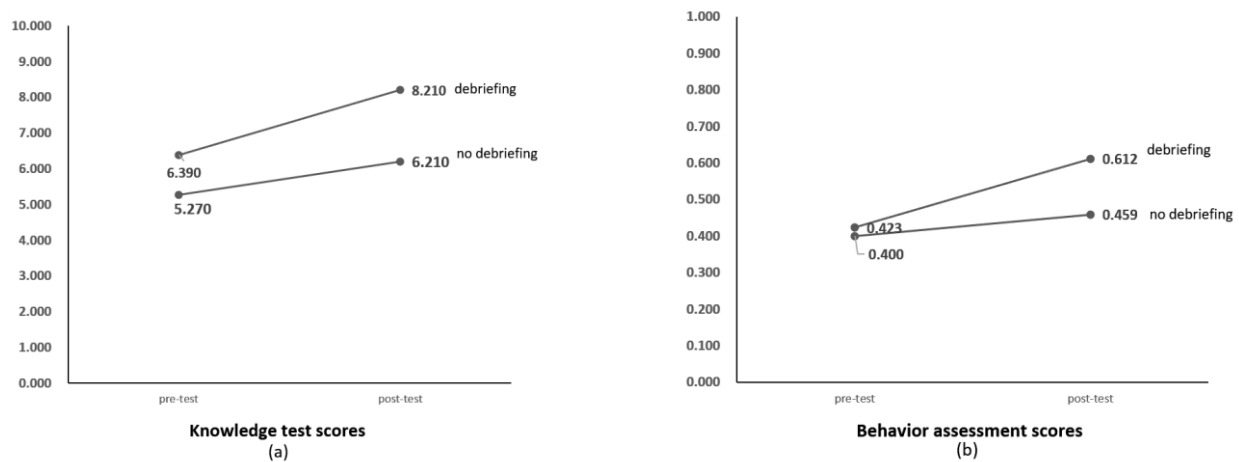

Figure 4. Comparison of the two groups of knowledge and behavior pretest and posttest scores. (**a**) improvement of knowledge test scores, (**b**) improvement of behavior assessment scores.

4.3. Variance in Learning Outcomes

To further explore the varying effect of debriefing on different types of knowledge and behavioral learning outcomes, independent sample *t*-tests were performed on the posttest sub-scores of both treatment and control groups. The group differences are plotted in Figure 5.
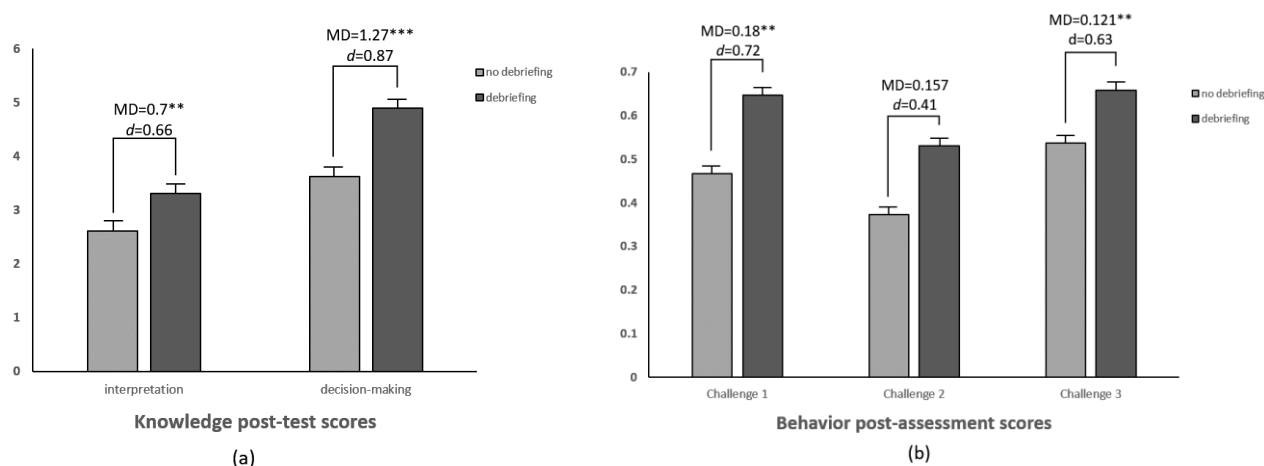
Figure 5. Comparison of the two groups' post-intervention assessment scores. (**a**) mean group differences in interpretation and decision-making scores in the knowledge test, (**b**) mean group differences in the behavior assessment scores from the three pedestrian challenges in IVR. ** The difference is significant at $p < 0.01$, *** The difference is significant at $p < 0.001$.

In terms of knowledge acquisition, Figure 5a shows that the treatment group showed significantly higher scores in both the interpretation and decision-making sections of the knowledge test ($MD_{(interpretation)} = 0.7$, SD = 0.24, $p = 0.005$; $MD_{(decision\ making)} = 1.27$, SD = 0.33, $p = 0.000$), in comparison with the control group. Further, the effect size of the mean difference in decision-making scores was considered large (Cohen's $d = 0.87$), while the effect size in interpretation scores was medium (Cohen's $d = 0.66$). The statistical results suggest that, despite its general effectiveness in promoting the participants' knowledge acquisition, the impact of debriefing on participants' decision making tended to be greater than that on interpretation.

In terms of behavioral performance, the treatment group received significantly higher behavior scores than the control group in two of the three challenges in the IVR environment, as seen in Figure 5b. The effect sizes of the mean differences for Challenge 1 and 3 were 0.72 and 0.63, respectively, which is considered a medium effect. The results indicated that, after participating in the debriefing session, the participants demonstrated improved pedestrian behaviors in two street-crossing risk scenarios themed by peer influence (C1) and backing vehicles (C3). Interestingly, despite the medium effect size of mean difference ($MD = 0.157$, $d = 0.41$), the discrepancy between the treatment and control group in Challenge 2 performance was statistically insignificant ($p = 0.079$). A plausible explanation is that the theme of a blinking traffic light in Challenge 2 was deemed more difficult by the participants, leading to greater standard deviation in the behavior scores of Challenge 2 (SD = 0.088) that that of the other two challenges (SD = 0.057, SD = 0.043). This result suggests that the difficulty of learning tasks merits our attention when examining the effectiveness of debriefing on IVR learning.

5. Discussion and Conclusions

Based on an IVR pedestrian safety education program, this study employed a randomized controlled trial to investigate the effects of a debriefing strategy on promoting knowledge acquisition and behavior performance in the IVR environment. The statistical results support the overall effectiveness of IVR as an educational technology and reveal the moderating effects of debriefing as an essential teaching strategy for IVR-based instruction. In sum, there are four main findings of the present study, which are discussed below.

First, the within-subjects effects confirm that the IVR learning environment can effectively improve children's pedestrian knowledge and performance. This finding is consistent with the previous literature [31–33] that supports the effectiveness of IVR in promoting experiential learning through a technology-mediated presence. The unique affordances of

IVR such as immersion, interaction, and safety are known to create high-fidelity simulation that is authentic, hazard-proof, and cost-effective [9–11].

Second, the results support the essential role of debriefing for IVR-based instruction due to its moderating effect on learning outcomes. A possible reason is that the debriefing activity provided students with an opportunity to assume an active role in both cognitive and meta-cognitive processes, such as recalling, analyzing, and self-evaluating their own IVR experiences. This finding corroborates the existing literature on debriefing and experiential learning: a structured debriefing can assist learners to integrate declarative knowledge and procedural knowledge [34], improve critical thinking skills [35], and promote knowledge construction and transfer [36], thereby enabling learners to assimilate the concrete learning experiences into their cognitive structure [24].

Third, although debriefing had a positive effect on knowledge acquisition and behavior improvement in the IVR environment, the moderating effect was found to be significant on behavior improvement only. This result is consistent with the findings of Schwebel et al. (2014), which revealed no obvious correlation between knowledge acquisition and behavior performance in IVR-based instruction, with behavior improvement identified as the main learning outcome [26]. A plausible explanation is that the IVR environment is often used for practice-oriented pedagogy where the behavior improvement is likely the result of trial-and-error and reinforcement rather than rational cognitive decisions. Unfortunately, the debriefing strategy in this study cannot effectively address this shortcoming of IVR-based instruction: the diagnostic analysis of pedestrian performance in the debriefing session might not reveal all the misconceptions of traffic rules, as some may be hidden behind correct behaviors. As Ormrod (2016) argued, learners' behavioral performance and cognitive representations are not completely mapped during the learning process [25].

Lastly, we found that the magnitude of the debriefing effect on IVR learning outcomes varied with different types of knowledge content and performance tasks. This finding indicates the complexity of simulated learning, which is supported by the debriefing literature. For example, the systematic review of Lee et al. (2020) showed that the structured debriefing strategies in general lead to improved clinical knowledge acquisition, yet the effect of debriefing differs between clinical reasoning and clinical judgement [37]. Similarly, Chronister et al. (2012) found that medical students' resuscitation skills showed no significant improvement after debriefing because of the contrastively diverse performance on different resuscitating tasks such as ventricular tachycardia recognition, rescue breaths, cardiopulmonary resuscitation, and defibrillation shock delivery [38]. However, the rationale behind the variance in debriefing outcomes remains largely unknown, which calls for future research to systematically investigate the relationship between debriefing strategies and learning outcomes.

6. Implications for Practice

Three implications for applying IVR in educational practice can be drawn from the research findings. Foremostly, the unique affordances of IVR support risk-free trial-and-error practice in virtual scenarios of safety hazard, which holds great potential for reforming and innovating safety education. Moreover, instructor-guided debriefing is a convenient yet highly effective teaching strategy that can greatly improve the effectiveness of IVR learning. Finally, when designing debriefing strategies for IVR-based instruction, the varying impact of debriefing on knowledge and behavior learning outcomes should be carefully considered. Accordingly, we put forward the following suggestions for designing and implementing IVR-based instruction.

First, the IVR experience should be considered as only one phase of experiential learning rather than the whole learning process. While acknowledging the affordances of IVR, the learning activities outside the virtual environment should not be ignored, and the facilitating role of the teacher should be further emphasized. The IVR environment provides students with personalized and contextual learning experiences, while the face-to-face debriefing can better support real-time communication and natural interaction

between teachers and students, which is conducive to developing meaningful teaching and learning activities such as evaluation, diagnosis, feedback, and reflection. The two complement each other well in IVR-based instruction.

Second, the general process of "experience–debriefing–experience" is recommended for IVR-based instruction. The initial experience session allows free exploration in the virtual world when completing the learning tasks. This instructional session engages students in situational recognition and embodied experiences, and serves as an assessment of students' initial knowledge level and behavioral performance. In the debriefing session, an instructor can provide students with personalized facilitation based on their performance in the previous session. This session enables abstract conceptualization of concrete experiences, diagnosis of problematic behaviors, construction of meaning, and knowledge transfer. Finally, the second experience in the IVR environment not only consolidates reflective learning but also serves as a post-intervention test for simulation and debriefing experiences. The effectiveness of such an IVR instructional process has been verified in this study.

Third, in the process of debriefing, attention should still be paid to the teaching of the knowledge content. Systematic instruction and personalized guidance should be integrated into the debriefing session to promote knowledge acquisition in particular. This study found that the effect of debriefing on promoting knowledge learning was less satisfactory. Therefore, a focused content instruction activity should be introduced to the debriefing session: Instructors can use diverse media resources such as mind maps and micro-lectures to give focused explanations of key knowledge points, evaluate the mastery of knowledge, and ensure that students' improved behaviors are accompanied by the improvement of their knowledge structure.

7. Limitations and Future Research Recommendations

The present study has several limitations that suggest avenues for future research. First, all participants were conveniently selected from one elementary school in central China, which might undermine the external validity of the research findings with questionable sample representativeness. The regional differences between urban and rural students in entry-level knowledge and pedestrian experiences might also influence the IVR learning outcomes. Therefore, future research should include a diverse population of participants to boost the generalizability of the empirical findings. Second, the learning outcomes were assessed immediately after the debriefing intervention; thus, the impact of debriefing on knowledge retention and behavior sustentation remains largely unknown. Future research should explore the delayed effects of both IVR and debriefing experience on knowledge and behavior learning. Third, it should be noted that the effectiveness of the debriefing strategy in the IVR learning program might vary depending on the subject being taught and school level of implementation. As a result, the findings in this study should be further verified in other educational contexts. Finally, the IVR intervention in the present study places high demands on technology and human resources as it requires plenty of IVR equipment and facilitators for safety supervision and individual debriefing, which reduces the likelihood of the research findings being applied to large-scale educational practice. Future research should explore effective approaches to increase the accessibility and cost-effectiveness of IVR-based instruction for increased educational adoption and sustained impact.

Author Contributions: Conceptualization, H.L. and W.L.; Formal analysis, Q.F. and Y.C.; Funding acquisition, H.L.; Investigation, Q.F., Y.C. and J.Z.; Supervision, H.L. and W.L.; Writing–original draft, Q.F.; Writing–review and editing, H.L. and W.L. All authors have read and agreed to the published version of the manuscript.

Appendix A. Child Pedestrian Behavior Observation List

Participant #:_____　　Time:_____　　Session #: pre-test

	Rubric	Challenge 1			Challenge 2			Challenge 3			Avg Score (Circled points / # of circles)
1	Attempt to dash into the street	Yes (0)	No (1)	-		-			-		
2	Attempt to cross when light is red	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	
3	Cross street when light is blinking (not waiting for the next signal)	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	
4	Check for traffic (left side) immediately before crossing	Yes (1)	No (0)	-	Yes (1)	No (0)	-	Yes (1)	No (0)	-	
5	Attempt to cross while bus backing (not waiting for bus to pass)		-			-		Yes (0)	No (1)	-	
	Avg Score (Circled points / # of circles)*										

* Do NOT count circled N/As.

Time:_____　　Session #: post-test

	Rubric	Challenge 1			Challenge 2			Challenge 3			Avg Score (Circled points / # of circles)
1	Attempt to dash into the street	Yes (0)	No (1)	-		-			-		
2	Attempt to cross when light is red	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	
3	Cross street when light is blinking (not waiting for the next signal)	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	Yes (0)	No (1)	N/A	
4	Check for traffic (left side) immediately before crossing	Yes (1)	No (0)	-	Yes (1)	No (0)	-	Yes (1)	No (0)	-	
5	Attempt to cross while bus backing (not waiting for bus to pass)		-			-		Yes (0)	No (1)	-	
	Avg Score (Circled points / # of circles)*										

* Do NOT count circled N/As.

References

1. Franzoni, V.; Milani, A.; Mengoni, P.; Piccinato, F. Artificial Intelligence Visual Metaphors in E-Learning Interfaces for Learning Analytics. *Appl. Sci.* **2020**, *10*, 7195. [CrossRef]
2. Luo, H.; Li, G.; Feng, Q.; Yang, Y.; Zuo, M. Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *J. Comput. Assist. Learn.* **2021**, *37*, 1–15. [CrossRef]
3. Franzoni, V.; Di Marco, N.; Biondi, G.; Milani, A. Virtual Reality for Enhancement of Emotional Mindset in the First Lockdown of United Kingdom for the Covid-19 Pandemics. In *Topics in Cryptology—CT-RSA 2020*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2021; pp. 189–198.
4. Di Natale, A.F.; Repetto, C.; Riva, G.; Villani, D. Immersive virtual reality in k-12 and higher education: A 10-year systematic review of empirical research. *Br. J. Educ. Technol.* **2020**, *51*, 2006–2033. [CrossRef]
5. Xie, Y.; Ryder, L.; Chen, Y. Using Interactive Virtual Reality Tools in an Advanced Chinese Language Class: A Case Study. *TechTrends* **2019**, *63*, 251–259. [CrossRef]
6. Calvert, J.; Abadia, R. Impact of immersing university and high school students in educational linear narratives using virtual reality technology. *Comput. Educ.* **2020**, *159*, 104005. [CrossRef]

7. Huang, W. Examining the impact of head-mounted display virtual reality on the science self-efficacy of high schoolers. *Interact. Learn. Environ.* **2019**, 1–13. [CrossRef]
8. Taranilla, R.V.; Cózar-Gutiérrez, R.; González-Calero, J.A.; Cirugeda, I.L. Strolling through a city of the Roman Empire: An analysis of the potential of virtual reality to teach history in Primary Education. *Interact. Learn. Environ.* **2019**, 1–11. [CrossRef]
9. Çakiroğlu, Ü.; Gökoğlu, S. Development of fire safety behavioral skills via virtual reality. *Comput. Educ.* **2019**, *133*, 56–68. [CrossRef]
10. Zhang, H.; He, X.; Mitri, H. Fuzzy comprehensive evaluation of virtual reality mine safety training system. *Saf. Sci.* **2019**, *120*, 341–351. [CrossRef]
11. Luo, H.; Yang, T.; Kwon, S.; Zuo, M.; Li, W.; Choi, I. Using virtual reality to identify and modify risky pedestrian behaviors amongst Chinese children. *Traffic Inj. Prev.* **2020**, *21*, 1–6. [CrossRef]
12. Wu, B.; Yu, X.; Gu, X. Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *Br. J. Educ. Technol.* **2020**, *51*, 1991–2005. [CrossRef]
13. Gardner, R. Introduction to debriefing. *Semin. Perinatol.* **2013**, *37*, 166–174. [CrossRef]
14. Moldjord, C.; Hybertsen, I.D. Training reflective processes in military aircrews through holistic debriefing: The importance of facilitator skills and development of trust. *Int. J. Train. Dev.* **2015**, *19*, 287–300. [CrossRef]
15. Brown, D.K.; Wong, A.H.; Ahmed, R.A. Evaluation of simulation debriefing methods with interprofessional learning. *J. Interprofessional Care* **2018**, *32*, 779–781. [CrossRef]
16. Odo, D.M. An Investigation of the development of pre-service teacher assessment literacy through individualized tutoring and peer debriefing. *J. Inq. Action Educ.* **2016**, *7*, 31–61.
17. Mete, D.E. Learning from reflection: A case of a post-sojourn debriefing workshop with EFL university students. *J. Lang. Linguist. Stud.* **2019**, *15*, 1376–1394. [CrossRef]
18. Aucott, C.; Soni, A. Reflections on the use of Critical Incident Stress Debriefing in schools. *Educ. Psychol. Pract.* **2016**, *32*, 85–99. [CrossRef]
19. Schary, D.P.; Jenny, S.E.; Morrow, G.S.; Wozniak, T. Bringing Challenge Course Activities into the Classroom: Pedagogical Strengths, Obstacles, and Recommendations. *J. Outdoor Recreat. Educ. Leadersh.* **2018**, *10*, 238–252. [CrossRef]
20. Kolb, D.A. *Experiential Learning: Experience as the Source of Learning and Development Second Edition*; Pearson Education Limited: Saddle Hill, NJ, USA, 2015.
21. Thatcher, D.C. Promoting Learning through Games and Simulations. *Simul. Gaming* **1990**, *21*, 262–273. [CrossRef]
22. Lederman, L.C. *Intercultural Communication, Simulation, and the Facilitative Role of the Teacher*; Far West Laboratory for Education Research and Development: San Francisco, CA, USA, 1983.
23. Pearson, M.; Smith, D. Debriefing in experience-based learning. *Reflect. Turn. Exp. Into Learn.* **1986**, *16*, 55–72.
24. Lederman, L.C. Differences that Make a Difference: Intercultural Communication, Simulation, and the Debriefing Process. *Global Interdepend.* **1992**, *1992*, 333. [CrossRef]
25. Ormrod, J.E. *Human Learning*, 7th ed.; Pearson Education: London, UK, 2015.
26. Schwebel, D.C.; McClure, L.A. Training children in pedestrian safety: Distinguishing gains in knowledge from gains in safe behavior. *J. Prim. Prev.* **2014**, *35*, 151–162. [CrossRef] [PubMed]
27. Anderson, J.R. Acquisition of Cognitive Skill. *Psychol. Rev.* **1982**, *89*, 369–406. [CrossRef]
28. Haye, A.; Torres-Sahli, M. To feel is to know relations: James' concept of stream of thought and contemporary studies on procedural knowledge. *New Ideas Psychol.* **2017**, *46*, 46–55. [CrossRef]
29. Fung, F.M.; Choo, W.Y.; Ardisara, A.; Zimmermann, C.D.; Watts, S.; Koscielniak, T.; Blanc, E.; Coumoul, X.; Dumke, R. Applying a Virtual Reality Platform in Environmental Chemistry Education To Conduct a Field Trip to an Overseas Site. *J. Chem. Educ.* **2019**, *96*, 382–386. [CrossRef]
30. Schwebel, D.C.; Wu, Y.; Li, P.; Severson, J.; He, Y.; Xiang, H.; Hu, G. Evaluating smartphone-based virtual reality to im-prove Chinese schoolchildren's pedestrian safety: A nonrandomized trial. *J. Pediatric Psychol.* **2018**, *43*, 473–484. [CrossRef] [PubMed]
31. Han, I. Immersive virtual field trips in education: A mixed-methods study on elementary students' presence and perceived learning. *Br. J. Educ. Technol.* **2020**, *51*, 420–435. [CrossRef]
32. Chen, X.M.; Chen, Z.B.; Li, Y.; He, T.Y.; Hou, J.H.; Liu, S.; He, Y. ImmerTai: Immersive motion learning in VR environments. *J. Vis. Commun. Image Represent.* **2019**, *58*, 416–427. [CrossRef]
33. Buttussi, F.; Chittaro, L. Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1063–1076. [CrossRef]
34. Al sabei, S.D.; Lasater, K. Simulation debriefing for clinical judgment development: A concept analysis. *Nurse Educ. Today* **2016**, *45*, 42–47. [CrossRef]
35. Kikkawa, Y.; Mavin, T.J. A review of debriefing practices: Toward a framework for airline pilot debriefing. *Aviat. Psychol. Appl. Hum. Factors* **2017**, *7*, 42–54. [CrossRef]
36. Tutticci, N.; Ryan, M.; Coyer, F.; Lewis, P.A. Collaborative facilitation of debrief after high-fidelity simulation and its implications for reflective thinking: Student experiences. *Stud. High. Educ.* **2018**, *43*, 1654–1667. [CrossRef]
37. Lee, J.H.; Lee, H.; Kim, S.; Choi, M.; Kim, S.H. Debriefing methods and learning outcomes in simulation nursing education: A systematic review and meta-analysis. *Nurse Educ. Today* **2020**, *87*, 104345. [CrossRef] [PubMed]
38. Chronister, C.; Brown, D. Comparison of Simulation Debriefing Methods. *Clin. Simul. Nurs.* **2012**, *8*, e281–e288. [CrossRef]