Use Virtual Reality to Enhance Intercultural Sensitivity: A Randomised Parallel Longitudinal Study

Chen Li, *Member, IEEE*, Angel Lo Lo Kon, and Horace Ho Shing Ip

Abstract—Prior studies suggest that emotional empathy is one of the components of intercultural sensitivity - the affective dimension under the concept of intercultural communication competence. Based on existing theories and findings, this paper reports a randomised parallel longitudinal study investigating the use of virtual reality (VR) exposure to enhance intercultural sensitivity. A total of 80 participants (36 females and 44 males) joined the study and were included in the data analysis. The participants were randomly assigned to the VR group, the video group, and the control group. Their intercultural sensitivity was measured three times: one week before the exposure (T_1), right after the exposure (T_2), and three weeks after the exposure (T_3). The results suggested that (1) the intercultural sensitivity of the VR group was significantly enhanced in both within-subject comparisons and between-subject comparisons, (2) there were no significant differences in intercultural sensitivity between the VR group and the video group at T_2, but the VR group retained the enhancement better at T_3, and (3) the sense of presence and emotional empathy well predicted the change in intercultural sensitivity of the VR group. The results, together with the participants' feedback and comments, provide new insights into the practice of using VR for intercultural sensitivity training and encourage future research on exploring the contributing factors of the results.

Index Terms—Virtual reality, presence, intercultural sensitivity, emotional empathy

1 Introduction

One of the most widely accepted definitions of intercultural sensitivity was proposed by Chen and Starosta under the concept of intercultural communication competence, which umbrellaed three dimensions - the cognitive dimension called cultural awareness, the behavioural dimension called intercultural adroitness, and lastly the affective dimension called intercultural sensitivity [14]. Chen and Starosta suggested that intercultural sensitivity mainly deals with our affects and is conceptually about how we could develop positive emotions towards "understanding and appreciating cultural differences that promotes an appropriate and effective behavior in intercultural communication" (p. 6) [14] . In today's global society, intercultural sensitivity is becoming increasingly important in various social contexts [34]. Hence, effectively enhancing intercultural sensitivity, especially among youth, is crucial for maintaining social solidarity and positive social relationships.

Empathy is believed to be one of the fundamental components contributing to intercultural sensitivity [7,8,15]. Although there are various definitions of empathy, they share the common idea of one person's responses, which could be cognitive, emotional, and/or behavioural, to his or her perceptions of another person's current situation or experience. Many believe empathy is the precursor to prosocial behaviours and can also act as a preventive factor against aggression [20, 23, 48]. In recent years, there has been a growing number of studies reporting the use of interactive media, such as video games and virtual reality (VR) experiences, to promote empathy and to elicit prosocial behaviours [26, 27, 46, 47]. Some practitioners even regard VR as the ultimate empathy machine due to its unique capabilities to induce perceptual illusions; it can evoke strong illusions of being in the virtual world and temporarily inhabiting other people of different ages, from different ethnic groups, of various social-economic statuses, or with special needs [24, 35, 44, 53, 54]. These perceptual illusions provide a realistic and immersive way of experiencing others' situations and making inferences about others' feelings and intentions given certain contexts; the ability to make such inferences is often referred to as cognitive empathy [28], while the vicarious emotional response to the perceived feeling of others is often referred to as emotional empathy or affective empathy [41]. A prior study suggests that VR is effective at enhancing emotional empathy but not cognitive empathy [40].

Given the potential of VR to enhance emotional empathy [40] and the connection between emotional empathy and intercultural sensitivity [7, 8, 15], there is a growing interest in using VR to enhance intercultural sensitivity. Meanwhile, there are new research questions to be addressed and knowledge gaps to be filled [1, 70]. For example, more rigorous evidence is needed to examine the effectiveness of using VR for intercultural sensitivity training; this requires the adoption of randomised controlled trials (RCTs) on larger samples. Additionally, the effectiveness should be compared to that of other media, such as video clips. Moreover, how the enhancement of intercultural sensitivity can be retained after exposure is largely underexplored. Lastly, the factors that could potentially contribute to effectiveness have not been examined empirically.

To address these new research questions and fill the knowledge gaps, a randomised parallel longitudinal study was conducted to empirically study the effectiveness of using VR for intercultural sensitivity training. A total of 80 participants joined the four-week-long study and were randomly assigned to the VR group, the video group, or the control group, adopting a 3 (groups) $\times$ 3 (time points) experimental design. Both the within-subject and between-subject comparisons demonstrated the effectiveness of using VR to enhance intercultural sensitivity; the VR group also retained the enhancement better than the video group. Moreover, the sense of presence and emotional empathy were found to be strong predictors of the enhancement of intercultural sensitivity among participants in the VR group. The results provided new insights into VR-enabled intercultural sensitivity training, as well as encouraged future research on exploring the contributing factors of the results. The major contributions of this study are as follows: (1) the use of VR to enhance intercultural sensitivity is strongly supported by the empirical evidence, and (2) future design of VR applications for enhancing intercultural sensitivity should aim to provide a strong sense of presence and to consider target audiences' emotional empathy.

- *Chen Li is with Department of Computing, The Hong Kong Polytechnic University. E-mail: richard-chen.li@polyu.edu.hk.*
- *Angel Lo Lo Kon is with Centre for Innovative Applications of Internet and Multimedia Technologies, City University of Hong Kong. E-mail: angelluluk@gmail.com.*
- *Horace Ho Shing Ip is with the Department of Computer Science and Centre for Innovative Applications of Internet and Multimedia Technologies, City University of Hong Kong. E-mail: horace.ip@cityu.edu.hk.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

2 Related Work

2.1 Intercultural Sensitivity and Empathy

Intercultural sensitivity is a dynamic concept, the definition of which has been evolving over recent decades. It has frequently been defined as an individual's awareness of cultural differences and the views of people from different cultures. Bennett adopted a developmental approach to describe intercultural sensitivity as a process of transforming oneself from ethnocentric stages (i.e., denial, defence, and minimisation) to ethnorelative stages (i.e., acceptance, adaptation, and integration) [8]; the transformation involves the affective, cognitive and behavioural aspects of intercultural interaction. Bhawuk and Brislin suggested that intercultural sensitivity should be associated with one's individualism and collectivism, flexibility, and open-mindedness [10]. Intercultural sensitivity involves various qualities; one should, for example, "be interested in other cultures, be sensitive enough to notice cultural differences, and then also be willing to modify their behaviour as an indication of respect for the people of other cultures" (p. 416) [10]. In measuring intercultural sensitivity, the authors mentioned the importance of evaluating one's ability to empathise with people from other cultures when they behave differently, and whether they can adjust their behaviour accordingly. Moreover, Bennett also highlighted the importance of empathy as the underlying mechanism of the adaptation stage, which enables individuals to shift between cultural contexts and adapt to cultural differences [6].

Chen and Starosta placed a greater emphasis on the affective side of intercultural competence. They defined intercultural sensitivity as the affective component in intercultural communication competence, a mindset that allows individuals to recognise the behavioural, perceptual, and emotional differences in their counterparts from different cultures [15]. Further, it is "an individual's ability to develop a positive emotion towards understanding and appreciating cultural differences that promotes an appropriate and effective behaviour in intercultural communication" (p. 5), which should be differentiated from the cognitive aspect (i.e., intercultural awareness) and the behavioural aspect (i.e., intercultural competence) [13]. According to this definition, intercultural sensitivity involves (1) self-esteem, which enables an individual to handle potential negative feelings such as alienation, frustration and stress brought by potential ambiguous situations in intercultural communication, (2) self-monitoring, which allows an individual to detect situational cues and behave appropriately, (3) open-mindedness, such that the individual can recognise, accept, and appreciate intercultural differences, (4) empathy, or being able to step into others' minds to understand their thoughts with vicarious feelings during interaction, (5) interaction involvement, or the ability to be responsive, attentive, and perceptive during interaction, and (6) non-judgement, or being sincere in listening to culturally different counterparts without jumping to conclusions.

As intercultural sensitivity is crucial to effective intercultural communication, researchers have investigated the effectiveness of various training approaches to enhance intercultural sensitivity (e.g., via movies [30] and creative imagination [21]). However, there are limited studies examining the effectiveness of using immersive media to enhance intercultural sensitivity. Meanwhile, it is evident that media with higher levels of interactivity and immersion can increase empathy [68]. In particular, a meta-analysis showed that VR is effective at enhancing emotional empathy but not cognitive empathy [40]. The authors suggested that VR content often did not require users to engage in effortful mental processing, which is essential for improving cognitive empathy, while on the other hand, emotional empathy involves a more automatic and less conscious mental process. As such, VR content can be more effective in eliciting emotional responses rather than encouraging users to engage in perspective taking. Thus, this study adopts the definition provided by Chen and Starosta regarding intercultural sensitivity because of their more substantial emphasis on the affective side of intercultural communication competence [15].

2.2 Presence

The term VR was first coined in the 1980s by Jaron Lanier to describe a computer-generated virtual world. The early definition and development of VR mainly focussed on enabling technologies, such as real-time motion tracking technologies to enable more intuitive interaction, computer graphics techniques and optimisations to support more realistic rendering of computer-generated virtual world, and display technologies to enable close-to-real-life visual perceptions [11]. Steuer suggested that "the key to defining VR in terms of human experience rather than technological hardware is the concept of presence" (p. 75) and that defining VR solely by technological capabilities lacked insights into the processes or effects of using VR [60]. The concept of presence, which might be interpreted slightly differently in various contexts, is now widely regarded as the desired design outcome of VR experiences and thus a metric for measuring the quality of the experiences. It is also considered a perceptual illusion that is quite unique to VR exposure [55]. Slater further developed the concept of presence by introducing place illusion (PI) and plausibility illusion (Psi), which are hypothesised to lead to the participants' realistic behaviours in virtual environments [56].

A few well-verified instruments have been developed to measure presence, such as the Igroup Presence Questionnaire (IPQ) [50], the Slater-Usoh-Steed (SUS) presence questionnaire [64], and the Witmer and Singer (WS) presence questionnaire [69]. Different instruments may adopt slightly different definitions of presence, thus, direct comparisons across different instruments could be unconscientious; in practice, these well-verified instruments can measure presence reliably given the corresponding definitions and contexts of presence.

Although it is unclear whether presence could have significant effects on VR-enabled intercultural sensitivity training, a few empirical studies have been carried out to study its relationship with empathy. Prior studies suggest that a high level of presence experienced by participants can be related to increased emotional reactivities [25, 67]. Bertrand et al. found that such emotional reactivities, which attracted their participants' full attention to the immersive and interactive storytelling, could improve the participants' empathy [9]. Shin drew a similar conclusion [53]; it was found that situational empathy mediated the effects of presence on engagement among participants with high levels of dispositional empathy when the contents were delivered through VR headsets.

2.3 Enhance Intercultural Sensitivity

Previous research studies have suggested that being exposed to intercultural environments could enhance intercultural sensitivity. For example, a great number of studies have investigated and observed the effects of studying abroad or attending international schools on students' intercultural sensitivity (e.g., [45,49,61,63]). In contrast to long-term exposure, the effects of relatively short exposure to intercultural environments on intercultural sensitivity have not been thoroughly investigated. Given the unique capability of VR that allows people to illusively experience situations from others' points of views [46], there are a few empirical studies that have investigated the effectiveness of VR in enhancing intercultural sensitivity or intercultural communication competencies in general. For example, Banakou et al. studied whether the virtual embodiment of white people in a black virtual body was effective in reducing implicit racial bias [4]. The results showed that implicit racial bias decreased more for those with the black virtual body than the white. Li et al. focussed on the sense of presence; by recreating landmark locations that ethnic minorities in Hong Kong often visited and using head-mounted displays (HMDs) to deliver the immersive experience, the study aimed to investigate how VR could change one's attitudes towards ethnic minorities [35]. The study found a statistically significant increase in intercultural sensitivity among 67 young Chinese participants after the VR exposure. Coffey et al. studied the efficacy of VR in intercultural sensitivity acquisition among undergraduate students [17]. The study found significant gains in intercultural sensitivity among the participants, who were exposed to the immersive experience built

with SecondLife™[1]. The gains were also found to be higher than for those who were exposed to the non-immersive web-based experience. Based on these empirical studies, it is believed that VR can be effective at enhancing intercultural sensitivity. However, whether and how the enhancement could be retained after the exposure is largely underexplored. Moreover, the factors that could potentially contribute to the effectiveness have not been examined empirically.

3 Research Questions

Based on prior research, this study aims to answer five major research questions.

RQ1. Can the intercultural sensitivity of the VR group be enhanced immediately after the exposure?

RQ2. Can the enhancement of the intercultural sensitivity of the VR group be retained three weeks after the exposure?

RQ3. Compared to the video group, how is the intercultural sensitivity of the VR group enhanced immediately after the exposure?

RQ4. Compared to the video group, how is the enhancement of the intercultural sensitivity of the VR group retained three weeks after the exposure?

RQ5. What are the predictive factors of the change in the intercultural sensitivity of the VR group?

Although enhancing intercultural sensitivity is an interesting application of VR for both practitioners and researchers, there is a lack of solid evidence on its effectiveness, especially when the exposure only occurs once. Additionally, how the change in intercultural sensitivity can be retained after exposure needs to be further explored via longitudinal studies. By answering **RQ1** and **RQ2**, this paper can fill these knowledge gaps.

At the same time, it is also important to compare the effectiveness of VR on intercultural sensitivity to the effectiveness of other media. This study also aims to address such concerns and test our hypotheses by answering **RQ3** and **RQ4**.

Lastly, there is not much work on exploring what factors could potentially contribute to the effectiveness of using VR for intercultural sensitivity training. The designing of the VR experience and scenarios is an expensive trial-and-error process. If the factors can be identified by answering **RQ5**, the design process can be more cost-effective. It is not feasible to investigate all potential factors in one study. Here, we focus on two potential factors - the sense of presence and the participants' emotional empathy - based on existing theories and findings. The sense of presence is a perceptual illusion during VR exposure [11, 56, 60]; as shown in previous surveys, VR is effective at enhancing empathy, especially emotional empathy [40, 66, 68], and the effectiveness is often hypothesised to be attributed to various perceptual illusions, including the sense of presence. Given the close relationship between intercultural sensitivity and emotional empathy, we chose to measure and analyse the sense of presence during VR exposure. The participants' emotional empathy was also measured and analysed in this study because prior studies also suggested that there were interactions between emotional empathy, presence, and intercultural sensitivity during and after the consumption of interactive media [2, 53].

Besides the five major research questions focussing on intercultural sensitivity, which is an affective outcome, we also want to explore the cognitive outcomes of the exposure (i.e., intercultural knowledge) and compare them with the affective outcomes as the secondary objectives of this study.

4 Method

4.1 Experimental Design

This longitudinal study employed a three-group parallel design spanning three time points over four weeks. Specifically, the three groups were the VR group, the video group, and the control group. The three time points are named T_1, T_2, and T_3 for the rest of the paper; T_1 was one week before the exposure, T_2 was right after the exposure, and T_3 was three weeks after the exposure. The participants' intercultural sensitivity and intercultural knowledge were measured at all three time points, emotional empathy as a trait was measured at T_1, and presence was measured at T_2 among participants in the VR group only. The measures taken at T_1 could be considered as a baseline to minimise any potential bias or imbalance of participants across the three groups, in addition to the randomisation process.

The experimental design can help us answer our research questions. Specifically, **RQ1** and **RQ2** can be answered by comparing the change in intercultural sensitivity between the VR group and the control group across three time points. **RQ3** and **RQ4** can be answered by comparing the change of intercultural sensitivity between the VR group and the video group across the three time points. The measures of emotional empathy and presence serve as the independent variables of the multiple linear regression models to help us answer **RQ5** (see Section 4.6).

Table 1. Demographics of sample participants.

Condition	Gender (male)	Ethnicity (minorities)	Age (years)
	n	n	mean (SD)
VR ($n = 27$)	12	4	22.03 (2.37)
Video ($n = 26$)	11	3	21.69 (2.11)
Control ($n = 27$)	13	2	22.48 (2.66)
Combined ($n = 80$)	36	9	22.07 (2.39)

4.2 Participants

An *a priori* power analysis was conducted using G*Power[2] to compute the needed sample size before the recruitment of the participants. The effect size in this study was set to 0.5, which is considered medium using Cohen's criteria [18]. To study the interaction effects using the 3 (groups) $\times$ 3 (time points) mixed analysis of variance (ANOVA), the total sample size needed was 78 when the error probability $alpha$ and power were set to 0.05 and 0.95, respectively [22].

The application for ethical review of this study was approved by the institutional review board of The Hong Kong Polytechnic University. A total of 90 individuals registered for participation through posted flyers on the university's campus. Four participants withdrew from the study due to time constraints. Five participants were excluded from the data analyses as they could not continue their participation at T_3 due to a sudden spike in coronavirus cases in Hong Kong and the strict social distancing rules imposed; and one participant in the video group was excluded from the data analyses due to lost contact after T_2 (see Figure 1). Hence, a total of 80 participants were included in our analyses. All of them had normal or correct-to-normal vision. None of them reported health conditions (e.g., epilepsy) that might cause safety or health concerns. Of the 80 participants included in our analyses, 71 participants identified themselves as the major ethnic group (i.e., Chinese) in Hong Kong, and the rest (11.25%) were ethnic minorities, including four from South Asia, three from West and Central Europe, one from North America, and one from Oceania. According to the government's latest census data published in 2016 [12], 8% of the resident population belonged to ethnic minorities. Hence, the ethnicity of the sample of this study was slightly more diverse than that of the resident population in Hong Kong. The demographics of the participants are presented in Table 1.

4.3 Procedure

Each registered participant was first assessed for eligibility. Eligible participants were randomly assigned to one of the three groups via a simple random number generator[3] and then invited to our laboratory at T_1. Upon their arrival, they were briefed about the procedure and

[1] https://secondlife.com/

[2] https://stats.idre.ucla.edu/other/gpower/

[3] https://www.randomizer.org/

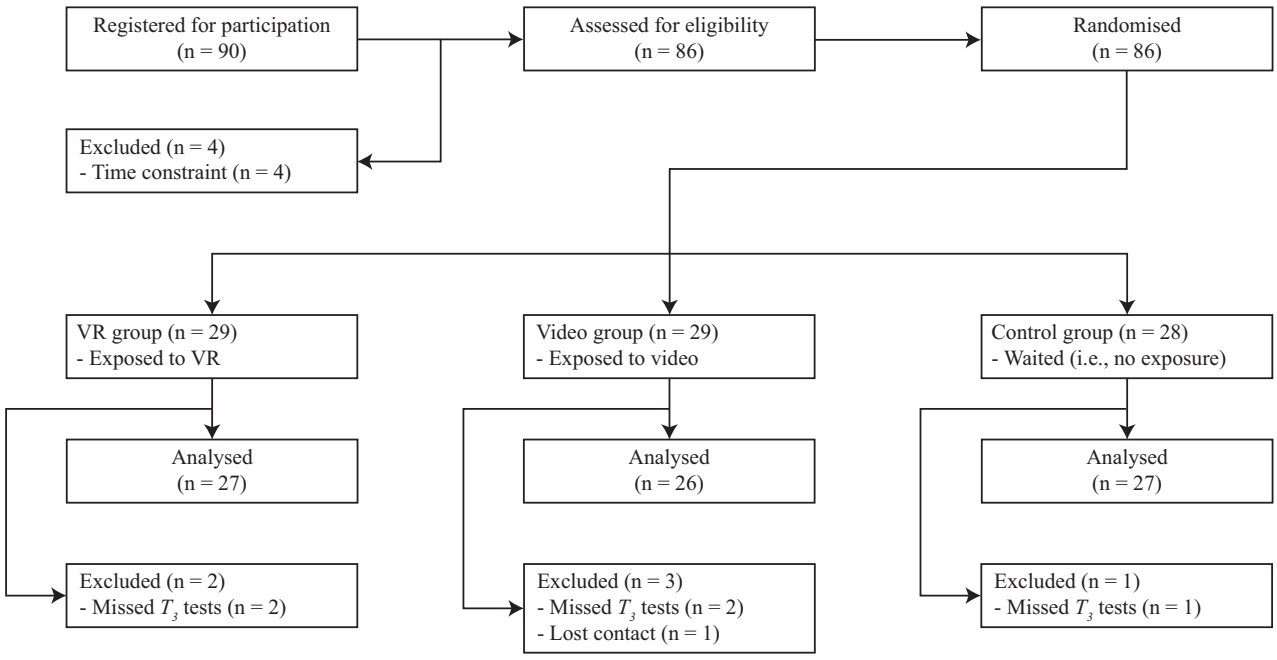

Fig. 1. The flow diagram of the experiment.

then read and signed a consent form. Participants were then asked to complete the demographics questionnaire, the Intercultural Sensitivity Scale (ISS) [15], the Empathic Concern subscale of the Interpersonal Reactivity Index (IRI-EC) [19], and the Intercultural Knowledge Test (IKT) (see Section 4.5). All participants were invited to come back to the laboratory in one week to continue their participation.

For the VR group, participants were invited to a one-hour VR exposure session at T_2 in our laboratory. Upon their arrival, they were briefed about the procedure and the safety arrangements. They were then instructed to stand at the centre of a 3 meters $\times$ 3 meters area and put on the HMD. The experimenter helped the participants to adjust the HMD to ensure they could see the content, especially the texts, clearly and comfortably. After that, participants started experiencing the VR scenarios (see Section 4.4). The function and operation of the trigger buttons on the controllers were first explained in the VR scenario. Participants needed to know how to interact with the virtual objects before they could proceed further. After experiencing all the VR scenarios, participants were asked to take off the HMD and complete the IPQ [50]. They were then asked to complete other questionnaires that were expected to be completed at T_2, including the ISS and the IKT. After completing the questionnaires, the experimenter interviewed the participants by asking one open-ended question - "what do you think of the VR scenarios you have just experienced?" The experimenter asked the participants to elaborate further if the answers were not clear or comprehensive enough. After completing the interview, the participants were told that the experiment session had finished and were invited to come back to the laboratory in three weeks (at T_3) to continue their participation.

Participants of the video group were also invited to a one-hour exposure session at T_2 in our laboratory. Upon their arrival, they were briefed about the procedure and then asked to sit in front of the computer monitor. The experimenter then played the video clips on the monitor, which were direct conversions of the VR scenarios (see Section 4.4). After watching the three video clips, participants were asked to complete the questionnaires that were expected to be completed at T_2, including the ISS and the IKT. After completing the questionnaires, participants were told that the experiment session had finished and were invited to return to the laboratory at T_3 to continue their participation.

Participants of the control group were invited to our laboratory to complete the questionnaires that were expected to be completed at T_2, including the ISS and the IKT. After completing the questionnaires, participants did not participate in any other activities and were invited to come back to the laboratory at T_3 to continue their participation.

Participants from all three groups were invited to come back to the laboratory at T_3 and were asked to complete the ISS and the IKT again on the same computer. Then, they were told that their participation had been completed and was very much appreciated. Supermarket coupons worth HK$100 (equivalent to US$12.8) were given to the participants.

During the entire process, the real objectives of the study were not revealed to any participants. The experimenter followed pre-drafted scripts when communicating with the participants. The scripts were carefully designed so that the real objectives could be concealed. The participants only knew they were invited to help "test a VR application" the research team had developed.

4.4 Materials and Apparatus

The materials used for this study are mainly the VR scenarios for the VR group and the converted video clips for the video group.

Three VR scenarios were created based on three typical locations reflecting the daily lives of ethnic minorities in Hong Kong. The three locations were associated with three topics, including family and history, religious beliefs and identity, and communal life in Hong Kong (see Figure 2). Using gamification techniques, discoverable spots and items were embedded in the scenarios like a classic puzzle game. Participants needed to navigate the scenarios, and once they were close to the spots and items, visual cues would light up and aid them in interacting with the discoverables and learning intercultural knowledge (see Figure 3). To make the objective clearer, a number was shown in the VR to indicate how many spots and items were yet to be discovered in the current scenario. The number of the discoverables was seven in the first scenario, nine in the second scenario, and three in the third scenario. In addition to the discoverables, the participants also needed to pick at least six questions they would like to ask in the third scenario. Details of the VR scenario design are as follows. The first scenario was reminiscent of typical households of ethnic minorities in Hong Kong; by virtually turning on the television in the scene, the participants were able to learn that one of the most common ways for ethnic minorities in Hong Kong to learn Cantonese is through watching local television programmes. In the second scenario, a virtual recreation of the Kowloon Mosque and Islamic Centre built using the photogrammetry method,

Fig. 2. The three VR scenarios that are captured as panoramic images.

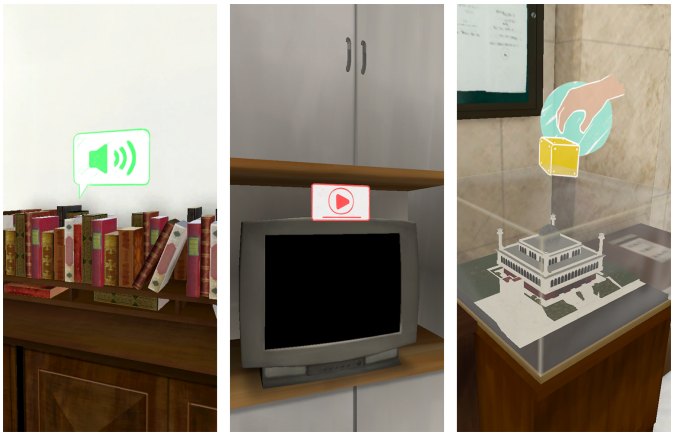

Fig. 3. The most common types of interesting spots and items embedded in the scenarios with corresponding visual cues; picture on the left - this item has corresponding voice narratives; picture in the middle - properly interact with this item to watch a video clip embedded in the scenario; picture on the right - this item can be virtually picked up for close examining.

the history of Muslims in Hong Kong was presented to the participants by moving close to and virtually touching a few historical pictures and the miniature of the building. In the third scenario, the participants were able to listen to recorded interviews of three ethnic minorities who moved from different areas of the world to Hong Kong more than a decade ago while being immersed in a community where both the Chinese and ethnic minorities live together. Participants in the VR group were able to experience the three VR scenarios at different paces. According to the system logs, the average exposure time of the VR group was 13 minutes 43 seconds, with the standard deviation as 57 seconds. OBS studio[4] was used to video record the screen while the experimenter experienced the three scenarios at an average pace. The resolution of the video clips was 2560×1440 at 60 frames per second. The three video clips corresponding to the three VR scenarios were 3 minutes 13 seconds, 5 minutes 12 seconds, and 4 minutes 6 seconds long, respectively; the total exposure time was 12 minutes 31 seconds. Considering the extra loading time (around 40 seconds in total) during the VR scenario switching in the runtime, the actual exposure time for both the VR group and the video group was similar.

The VR group experienced the scenarios using the Oculus Quest 2 HMD, which was tethered to a desktop computer via the Oculus Air Link wireless connection. The audio of the VR scenarios was played through the HMD's built-in speakers. The desktop computer was equipped with an Intel Core i7-10700K processor, 32GB RAM, and NVIDIA RTX 3080 graphics card. Using the software development and debugging tools provided by Oculus Integration SDK for Unity[5], it was confirmed that during the entire exposure, the HMD was constantly working at its 90Hz maximum refresh rate. The interactions were enabled by the two Oculus Touch controllers. A 3 meters $\times$ 3 meters area was partitioned in our laboratory for the experiment. Since the scenarios were created upon three-dimensional models of the locations, six-degree-of-freedom (6-DoF) movement was supported, which facilitated the sensorimotor contingencies and, in turn, should have invoked a high sense of presence during exposure [57]. The classical teleportation method was adopted to let the participants navigate the VR scenarios for considerable distances, although participants could still physically walk around within the partitioned area for shorter distances. The videos were delivered on a 27-inch monitor with built-in speakers connected to the same computer. The resolution and refresh rate of the monitor was set to 2560×1440 and 60Hz, respectively. We chose to use the monitor's built-in speakers instead of headphones because the exposure took place in our laboratory with good sound isolation, and the VR group also used the HMD's built-in speakers instead of headphones. We think the difference in sound quality between the monitor's built-in speakers and the HMD's built-in speakers was minimal and can be ignored, although we did not have the necessary equipment to measure the difference objectively.

[4]https://obsproject.com/

[5]https://developer.oculus.com/downloads/package/unity-integration/

4.5 Measures

4.5.1 Intercultural Sensitivity

The Intercultural Sensitivity Scale (ISS) [15] was adopted to measure participants' intercultural sensitivity. It comprises five subscales with a total of 24 items. The subscales are interaction engagement, respect for cultural differences, interaction confidence, interaction enjoyment, and interaction attentiveness. Responses are given on a five-point Likert scale: "strongly disagree" (1), "disagree" (2), "uncertain" (3), "agree" (4), and "strongly agree" (5). ISS has been used in studies from different cultures, including Chinese [35]. The reliability of the current sample is 0.731 (Cronbach's α) at T_1, 0.775 at T_2, and 0.799 at T_3.

4.5.2 Emotional Empathy

The Empathic Concern subscale of the Interpersonal Reactivity Index (IRI-EC) [19] was used to measure the participants' emotional empathy. The subscale has a total of seven items on a five-point Likert scale ranging from "Does not describe me well" (1) to "Describes me very well" (5). The reliability of the current sample is 0.712 (Cronbach's α).

4.5.3 Presence

The Igroup Presence Questionnaire (IPQ) [50] was used to assess the sense of presence among the participants of the VR group. The questionnaire has 14 items measured on a five-point scale, with higher scores indicating higher presence. Some items were adopted from previously published scales. The original wordings and anchors were kept, so the wordings of the response options vary across the items. For instance, one item states, "In the computer generated world I had a sense of 'being there'" which is adopted from [59], and the corresponding response anchors are between "not at all" (1) and "very much" (5). The reliability of the current sample is 0.855 (Cronbach's α).

4.5.4 Intercultural Knowledge Test

A nine-item multiple-choice Intercultural Knowledge Test (IKT) was designed to assess the participants' intercultural knowledge. The correct

answers to the IKT items were conveyed in the VR scenarios; each of the three VR scenarios had three corresponding items in the IKT. As the videos used in the experiment were directly converted from the VR scenarios, the correct answers to the IKT items were also conveyed to the video group but not to the control group. Participants received one mark for choosing the correct answer for each item, and the marks were summed up to be the total score for data analysis; hence, the maximum score was nine.

4.6 Data Analysis

Mixed analysis of variance (ANOVA) was used to understand if there were any interaction effects between the within-subject factor (i.e., three time points) and the between-subject factor (i.e., three groups) on the ISS scores. To further examine the six simple main effects, pairwise comparisons with Bonferroni adjustment were performed. The techniques above were repeated with the IKT scores as the dependent variable. Linear regression models were used to investigate the potential predictors of the ISS scores. A significance level of 0.05 was used across all analyses. All data analyses were conducted using IBM SPSS Statistics 26[6].

5 Results

5.1 Descriptive Statistics and the Baseline Measures

The descriptive statistics of the measures are shown in Table 2. One-way ANOVA was first used to check if there were any differences in the baseline measures of emotional empathy, intercultural sensitivity, and intercultural knowledge across the three groups at T_1. The results suggested that across the three groups, there were no significant differences in the mean IRI-EC scores ($F(2,77) = 0.536$, $p = 0.587$, $\eta_p^2 = 0.014$), the mean ISS scores ($F(2,77) = 0.038$, $p = 0.962$, $\eta_p^2 = 0.001$), or the mean IKT scores ($F(2,77) = 0.310$, $p = 0.734$, $\eta_p^2 = 0.008$).

5.2 Changes in Intercultural Sensitivity

The changes in the mean ISS scores across the three groups are plotted in Figure 4.

A 3×3 mixed ANOVA was applied to understand if there were any interaction effects between time and group on the ISS scores. There was a significant interaction between time and group on ISS scores ($F(4,154) = 7.411$, $p < 0.001$, $\eta_p^2 = 0.161$). Regarding the within-subject effects, significant differences were also found across the three time points ($F(2,154) = 22.860$, $p < 0.001$, $\eta_p^2 = 0.229$). Regarding the between-subject effects, no significant difference were found across the three groups ($F(2,77) = 2.061$, $p = 0.134$, $\eta_p^2 = 0.051$).

Pairwise comparisons with Bonferroni adjustment were performed to further check the simple main effects. At T_2, the simple main effect of group was significant ($F(2,77) = 6.022$, $p = 0.004$, $\eta_p^2 = 0.135$). At T_3, the simple main effect of group was also significant ($F(2,77) = 3.602$, $p = 0.032$, $\eta_p^2 = 0.086$). The detailed pairwise comparison results are shown in Table 3. For the VR group, the simple main effect of time was significant ($F(2,52) = 20.788$, $p < 0.001$, $\eta_p^2 = 0.444$). For the video group, the simple main effect of time was also significant ($F(2,50) = 11.711$, $p < 0.001$, $\eta_p^2 = 0.319$). For the control group, the simple main effect of time was not significant ($F(2,52) = 0.007$, $p = 0.993$, $\eta_p^2 < 0.001$). The detailed pairwise comparison results are shown in Table 4.

5.3 Changes in Intercultural Knowledge

The changes in the mean IKT scores across the three groups are plotted in Figure 5.

A 3×3 mixed ANOVA was applied to understand if there was an interaction between time and group on the IKT scores. There was a significant interaction between time and group on IKT scores ($F(4,154) = 29.323$, $p < 0.001$, $\eta_p^2 = 0.432$). Regarding the within-subject effects, significant differences were also found across the three time points ($F(2,154) = 132.522$, $p < 0.001$, $\eta_p^2 = 0.632$). Regarding the between-subject effects, significant differences were found across the three groups as well ($F(2,77) = 24.880$, $p < 0.001$, $\eta_p^2 = 0.393$).

Pairwise comparisons with Bonferroni adjustment were performed to further check the simple main effects. At T_2, the simple main effect of groups was significant ($F(2,77) = 58.111$, $p < 0.001$, $\eta_p^2 = 0.601$). At T_3, the simple main effect of groups was also significant ($F(2,77) = 12.492$, $p < 0.001$, $\eta_p^2 = 0.245$). The detailed pairwise comparison results are shown in Table 5. For the VR group, the simple main effect of time was significant ($F(2,52) = 60.694$, $p < 0.001$, $\eta_p^2 = 0.700$). For the video group, the simple main effect of time was also significant ($F(2,50) = 175.000$, $p < 0.001$, $\eta_p^2 = 0.875$). For the control group, the simple main effect of time was not significant ($F(2,52) = 0.388$, $p = 0.680$, $\eta_p^2 = 0.015$). The detailed pairwise comparison results are shown in Table 6.

5.4 Predictors of the Change in Intercultural Sensitivity

The bivariate correlation test was first conducted to check the collinearity among the independent variables (i.e., the IRI-EC scores, the IPQ scores, and age measured in months at T_1 of the participants of the VR group). The results suggested that there were no significant correlations between any pair of these independent variables.

A hierarchical multiple linear regression was employed to explore the relations between the independent variables and the dependent variable (i.e., the change of ISS scores from T_1 to T_2). In the first block, the participants' ages and genders were controlled, and they did not significantly contribute to the dependent variable ($F(2,24) = 1.201$, $p = 0.318$). In the second block, when the IRI-EC and IPQ scores were entered into the regression model as predictors, the model was statistically significant ($F(4,22) = 9.805$, $p < 0.001$), and 64.07% (R^2) of the variation in the dependent variable was explained. The standardised regression coefficient (β) of the IRI-EC scores was 0.494 ($p = 0.002$); the standardised regression coefficient (β) of the IPQ scores was 0.432 ($p = 0.006$).

Similar techniques were used to explore whether the IRI-EC and IPQ scores of the participants of the VR group could also predict the change of ISS scores from T_1 to T_3. In the first block, the participants' ages and genders were controlled, and they did not significantly contribute to the dependent variable ($F(2,24) = 0.233$, $p = 0.794$). In the second block, when the IRI-EC and IPQ scores were entered into the regression model as predictors, the model was statistically significant ($F(4,22) = 8.680$, $p < 0.001$), and 61.21% (R^2) of the variation in the dependent variable was explained. The standardised regression coefficient (β) of the IRI-EC scores was 0.700 ($p < 0.001$); the standardised regression coefficient (β) of the IPQ scores was 0.174 ($p = 0.254$).

6 Discussion

6.1 The Enhancement of Intercultural Sensitivity

Regarding the enhancement of intercultural sensitivity, the 3×3 mixed ANOVA suggested a significant interaction between time and group. The pairwise comparisons with Bonferroni adjustment further helped us look into the simple main effects and answer our research questions. The first research question we wish to address through this study is whether the intercultural sensitivity of the VR group could be enhanced immediately after the exposure (**RQ1**). The between-subject results suggested that the mean ISS scores of the VR group at T_2 were significantly higher than the mean ISS scores of the control group. Additionally, the within-subject results found that the mean ISS scores of the VR group at T_2 were significantly higher than at T_1. These results demonstrated the effectiveness of using VR for intercultural sensitivity training. Moreover, the between-subject and within-subject evidence suggested that the enhancement of intercultural sensitivity of the VR group could be retained three weeks after the exposure (**RQ2**). Specifically, the mean ISS scores of the VR group at T_3 were still significantly higher than the mean ISS scores of the control group, and the mean ISS scores of the VR group at T_3 were not significantly lower than at T_2.

Next, the aforementioned data analyses also helped us to address **RQ3** and **RQ4**. Regarding **RQ3**, the between-subject comparisons suggested that both the VR group and the video group exhibited a

[6]https://www.ibm.com/products/spss-statistics

Table 2. Descriptive statistics of the IRI-EC scores at T_1, the IPQ scores at T_2 (of the VR group only), and the ISS and IKT scores at all three time points.

Group	IRI-EC	IPQ	ISS			IKT		
	T_1	T_2	T_1	T_2	T_3	T_1	T_2	T_3
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
VR (n=27)	19.89 (3.215)	3.56 (0.446)	3.78 (0.290)	3.99 (0.226)	3.95 (0.237)	3.15 (1.027)	7.30 (1.613)	4.85 (1.512)
Video (n=26)	18.96 (3.493)	-	3.79 (0.196)	3.96 (0.206)	3.84 (0.234)	2.88 (1.143)	6.58 (1.391)	4.65 (1.384)
Control (n=27)	19.37 (3.103)	-	3.80 (0.179)	3.80 (0.175)	3.81 (0.188)	3.04 (1.454)	3.22 (1.423)	3.04 (1.480)

Table 3. Pairwise comparisons with Bonferroni adjustment of groups on the ISS scores at T_2 and T_3. The simple main effects of group son the ISS scores at T_1 were not significant thus no pairwise comparisons were performed.

Time	Pair	Mean Diff. (SE)	p	95% CI
T_2	VR-Control	0.188 (0.058)	0.005	[0.047, 0.330]
	Video-Control	0.156 (0.058)	0.028	[0.013, 0.299]
	VR-Video	0.033 (0.058)	1.000	[-0.111, 0.176]
T_3	VR-Control	0.156 (0.060)	0.034	[0.009, 0.303]
	Video-Control	0.042 (0.061)	1.000	[-0.106, 0.190]
	VR-Video	0.114 (0.061)	0.192	[-0.034, 0.262]

Table 4. Pairwise comparisons with Bonferroni adjustment of time on the ISS scores of the VR and video group. The simple main effects of time on the ISS scores of the control group were not significant thus no pairwise comparisons were performed.

Group	Pair	Mean Diff. (SE)	p	95% CI
VR	T_3-T_1	0.174 (0.038)	<0.001	[0.081, 0.267]
	T_2-T_1	0.208 (0.028)	<0.001	[0.140, 0.277]
	T_3-T_2	-0.034 (0.030)	0.769	[-0.107, 0.039]
Video	T_3-T_1	0.050 (0.039)	0.609	[-0.045, 0.144]
	T_2-T_1	0.165 (0.028)	<0.001	[0.095, 0.235]
	T_3-T_2	-0.115 (0.030)	0.001	[-0.189, -0.041]

Table 5. Pairwise comparisons with Bonferroni adjustment of groups on the IKT scores at T_2 and T_3. The simple main effects of groups on the IKT scores at T_1 were not significant thus no pairwise comparisons were performed.

Time	Pair	Mean Diff (SE)	p	95% CI
T_2	VR-Control	4.074 (0.403)	<0.001	[3.088, 5.060]
	Video-Control	3.355 (0.407)	<0.001	[2.359, 4.350]
	VR-Video	0.719 (0.407)	0.243	[-0.276, 1.715]
T_3	VR-Control	1.815 (0.398)	<0.001	[0.842, 2.788]
	Video-Control	1.617 (0.401)	<0.001	[0.635, 2.599]
	VR-Video	0.198 (0.401)	1.000	[-0.784, 1.180]

Table 6. Pairwise comparisons with Bonferroni adjustment of time on the IKT scores of the VR and video group. The simple main effects of time on the IKT scores of the control group were not significant thus no pairwise comparisons were performed.

Group	Pair	Mean Diff. (SE)	p	95% CI
VR	T_3-T_1	1.704 (0.254)	<0.001	[1.082, 2.326]
	T_2-T_1	4.148 (0.298)	<0.001	[3.418, 4.878]
	T_3-T_2	-2.444 (0.296)	<0.001	[-3.170, -1.719]
Video	T_3-T_1	1.769 (0.259)	<0.001	[1.135, 2.403]
	T_2-T_1	3.692 (0.407)	<0.001	[2.948, 4.436]
	T_3-T_2	-1.923 (0.302)	<0.001	[-2.662, -1.184]

similar enhancement of intercultural sensitivity immediately after the exposure at T_2, but the mean ISS scores of the VR group were slightly higher than the mean ISS scores of the video group. Regarding **RQ4**, the between-subject evidence suggested that the VR group retained the enhancement better than the video group did at T_3, suggesting the advantages of VR over less immersive and less interactive media for enhancing intercultural sensitivity.

These results bring two major implications to the application of VR for intercultural sensitivity training. First, the strong evidence on the enhancement of intercultural sensitivity and retention encourages future practices of VR-enabled intercultural sensitivity training. Previously, although the enhancement right after the VR exposure had been found and reported in other publications (e.g., [1,35]), there was little evidence of the retention of the enhancement. Our results helped confirm the desired outcomes of using VR for intercultural sensitivity training. Additionally, given the current global pandemic, being exposed to delicately designed VR scenarios could be an effective alternative to conventional face-to-face means of intercultural sensitivity training, such as workshops and guided tours. Second, the differences found regarding the retained enhancement of intercultural sensitivity between the VR group and the video group exhibited the advantage of VR exposure - the VR group retained the enhancement of intercultural sensitivity better than the video group did at T_3. To our best knowledge, such differences had been rarely explored in previous studies and might help justify the relatively higher costs of making the VR scenarios and purchasing the HMDs. From the research perspective, the findings also encourage further investigations on the change in intercultural sensitivity that go beyond three weeks after the exposure.

6.2 The Enhancement of Intercultural Knowledge

The changes in intercultural knowledge also provided some interesting insights, especially when the results were compared across the three groups at the three time points. Specifically, at T_2, the pairwise comparisons suggested that there were no significant differences in the mean IKT scores between the VR group and the video group. The mean IKT scores of the VR group and the video group were both significantly higher than the mean IKT scores of the control group at T_2. There were no significant differences between the mean IKT scores of the VR group and the scores of the video group. Similarly, at T_3, the mean IKT scores of the VR group and the video group were both significantly higher than the mean IKT scores of the control group, but no significant differences in the mean IKT scores were found between the two groups. Hence, our results suggested that VR and video were similarly effective in enhancing and retaining the participants' intercultural knowledge, although the retention was poorer than expected. Because intercultural knowledge is more of a cognitive outcome, our findings generally aligned with prior studies in educational research, which suggested that various media, ranging from conventional media to more interactive and more immersive media, such as VR, might lead to similar cognitive outcomes [16, 39, 62].

Meanwhile, it is worth noting that the instrument we used to measure intercultural knowledge (i.e., IKT) was mainly to measure participants' ability to memorise and recall the facts conveyed in the VR scenarios; this should be considered a narrowed view of intercultural knowledge [14, 38]. There should be more cognitive outcomes of the exposure that were not directly or quantitatively measured in our study. However, the contrast between the emotional outcomes (i.e., intercultural sensitivity) and the cognitive outcomes is an interesting finding; it

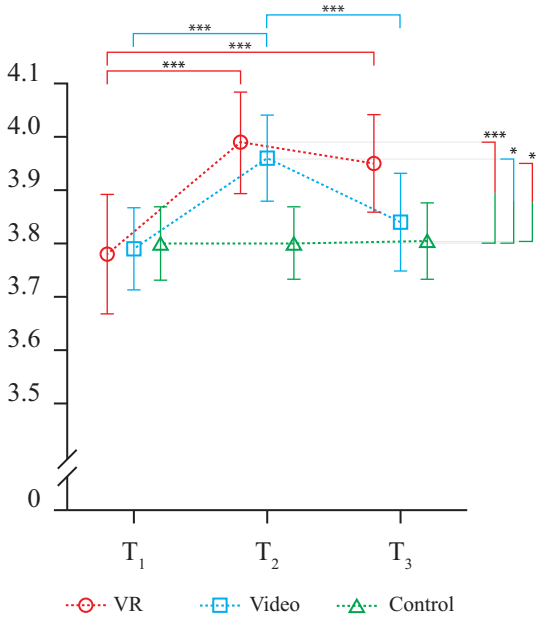

Fig. 4. The mean ISS scores of the three groups at three time points. The error bars indicate ±2 standard error.

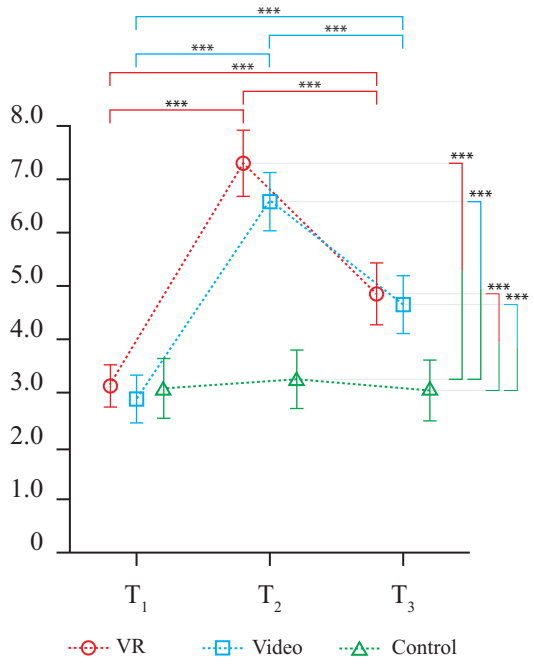

Fig. 5. The mean IKT scores of the three groups at three time points. The error bars indicate ±2 standard error.

encourages future studies to further investigate and compare the emotional outcomes to the cognitive outcomes in VR-enabled intercultural sensitivity training, as well as other application contexts.

6.3 The Predictors of the Enhancement

Based on existing findings and theories, this study particularly examined and investigated whether the sense of presence and the participants' emotional empathy could be predictive factors of the changes to the intercultural sensitivity of the VR group across the three time points (**RQ5**).

Using the linear regression models with IPQ and IRI-EC scores as the independent variables and the change of ISS scores between T_1 and T_2 as the dependent variable, our results suggested that both the sense of presence and the participants' emotional empathy could well predict the change of intercultural sensitivity between T_1 and T_2. A second linear regression model with the change of ISS scores between T_1 and T_3 suggested that the retention of intercultural sensitivity within three weeks after the exposure could also be well predicted by emotional empathy, but the sense of presence was a relatively weaker predictor in this case.

The findings here are exciting and bring a few important implications to practice as well as research. From the perspective of applying VR for intercultural sensitivity training, since both presence and emotional empathy were found to be predictors of the enhancement of intercultural sensitivity, the future design of using VR to enhance intercultural sensitivity can directly benefit from prior studies on how to better enhance the sense of presence during VR exposure (e.g., [36,50–52,55,56,58,65]). The approach can also potentially reduce the costs of making such VR applications. Meanwhile, from the research perspective, although the results are close to our expectations which were based upon prior theories, the linear regression models still could not fully explain the variations in the dependent variables, suggesting there are other factors that could also contribute to the change of intercultural sensitivity. Moreover, our results showed that age and gender had insignificant contributions to the enhancement of intercultural sensitivity. Hence, future research should measure other factors and include them in the predictive models.

6.4 Participants' Feedback and Comments

Participants in the VR group provided valuable feedback regarding their experiences and their impressions of the VR scenarios. Many of them used the terms "immersive", "engaging", and "realistic" to describe their experiences and the VR scenarios. Among those participants, three participants (participant A (female, 20 years old), B (male, 24 years old), and C (female, 19 years old)) said that they previously joined a tour of the Kowloon Mosque and Islamic Centre organised by a local non-governmental organisation; they all mentioned that they felt the venues in the VR scenario were extraordinarily similar to what they saw during the tour. Participant A said that she preferred the VR exposure over the tour because "[the information provided in] the VR scenario was more comprehensive and much easier to follow." Participant B said that after experiencing the VR scenario, he wanted to "revisit the location with friends." Participant C asked whether the community in the third VR scenario was real and where it was located. After knowing the location, she said that she grew up in the neighbourhood, but she moved to another district of the city over a decade ago. She had a strong feeling of the area when she was experiencing the scenario. Another participant (participant D (male, 21 years old)) who was studying computer science at the university asked the experimenter "how these very realistic scenes were made" and expressed a strong interest in learning the techniques. Such feedback and comments suggested that many of the participants indeed experienced high levels of presence when being exposed to the VR scenarios. This aligned with the quantitative data collected using the IPQ. The qualitative feedback also suggested a high level of engagement, which was not measured quantitatively in our study. Participants in the VR group also shared their thoughts regarding their personal experiences with intercultural sensitivity and intercultural communication when being asked to provide feedback and comments after the exposure. One participant (participant E (male, 21 years old)) said that he had "some prejudice towards ethnic minorities in Hong Kong" and tried to "avoid having conversations with them", but the VR scenarios helped him "eliminate the prejudice" and became "more willing to talk with them." Another participant (participant F (female, 24 years old)) said that she was often very curious about the cultures and religious beliefs of ethnic minorities in Hong Kong. Before joining the experiment, she often felt that "their cultures and [religious] beliefs are not approachable." But her thoughts and feelings had changed after seeing their willingness to share and discuss in the VR scenarios. These feedback and comments presented strong qualitative evidence on the effectiveness for VR exposure on enhancing intercultural sensitivity.

The interviews also allowed us to check if there were any potential improvements that could be made to the VR scenarios. Two participants (participants G (female, 21 years old) and H (male, 23 years old)) with no previous experience with VR said that they felt confused about teleportation in the VR scenarios. Participant G said that she had "never tried VR before" and found the teleportation technique "somewhat confusing"; by the end of the exposure, she still "felt lost and disoriented in the scenarios when jumping from one point to the other." Although teleportation is a widely used locomotion technique that can reduce VR sickness, this technique that discontinuously and instantly translates users from one location to the other, is known to induce spatial disorientation [3]. Participant H said that he was "confused at the beginning but luckily [the space of] the first scene was not large", so he could "learn how the [teleportation] technique works and better move and orient himself" in the second and third scenarios. Moreover, three participants (participants A, I (female, 27 years old), and J (male, 23 years old)) said that they wished the virtual characters in the VR scenarios could be more interactive, because except for the characters in the third scenario, most of the other characters could not be interacted with. Participant H said that he actually thought about whether a set of more interactive characters could be made so that he could practise intercultural communication with these virtual characters to avoid potential embarrassment in the real life. This could be an interesting and promising research direction; enabling technologies (e.g., natural language processing) can well support such training in VR with intelligent virtual agents [33, 37, 43].

7 Limitations

First, although our results provided both within-subject and between-subject evidence regarding the effectiveness of VR exposure on intercultural sensitivity and intercultural knowledge, the difficulty of recruiting participants during the pandemic, the dropouts due to a sudden spike in coronavirus cases in Hong Kong and the strict social distancing rules hindered us from collecting empirical evidence from more diverse subjects; the participants of the current sample had to be recruited using convenience sampling. This could hinder the generalisation of the study's results to other regions of the world because intercultural sensitivity training could be cultural, geographical, and religious dependent. Nevertheless, we hope that our study can serve as a solid foundation for future studies on more diverse samples.

Second, although the linear regression models fitted quite well, there were still variations in the changes in the VR group's intercultural sensitivity across the three time points that could not be fully explained using the models. As stated previously (see Section 3), it was impossible to include every potential factor in the models. Nevertheless, there are a few potential factors that need to be further examined in future studies. For example, the sense of embodiment might also predict changes in intercultural sensitivity before and after VR exposure. Although there are very few studies that investigate the relations between embodiment and intercultural sensitivity, prior studies have revealed the relations between embodiment and empathy (e.g., [5, 32, 53]). Kilteni et al. suggested that the sense of embodiment should be defined as a combination of three subcomponents, i.e., the sense of self-location, the sense of agency, and the sense of body ownership [31]. The use of self-avatars that are driven by users' body motions and facial expressions is a common way of invoking the sense of embodiment during VR exposure. But due to time constraints, this approach was not implemented in our VR scenarios. Hence, it was not feasible to further investigate whether the sense of embodiment could also predict changes in intercultural sensitivity. Other factors, such as user experience and engagement, could also interact with the changes in intercultural sensitivity. We will conduct follow-up studies to investigate these factors and the hypothesised interactions further.

Third, we only compared the effects of VR to conventional two-dimensional videos on intercultural sensitivity training. The participants in the video group were exposed to a less immersive and less interactive experience compared to those of the VR group. It would be interesting to further study the effects by manipulating the interactivity and immersion of the experience independently. We plan to conduct a further study by adding a non-interactive cinematic VR (i.e., 360-degree video) condition to the experiment design.

Finally, the novelty effect of VR was not thoroughly examined in this study. The novelty effect of VR on learning has been reported in previous studies (e.g., [29, 42]). However, whether VR has any novelty effect on intercultural sensitivity training is largely unknown. Because most of the participants had little VR experience before joining this study, it is plausible that the observed effect of VR on intercultural sensitivity was partly due to the novelty effect. To study it would require a new experiment design, which is beyond the scope of our study. We encourage future research to examine this topic and investigate whether the approach would become less effective when the exposure time to VR increases and participants become more familiar with the medium.

8 Conclusion

Following the definition of intercultural sensitivity and its close relation with emotional empathy, we designed this randomised parallel longitudinal study to investigate the use of VR exposure for enhancing intercultural sensitivity. It was found that the enhancement was significant and could be retained for at least three weeks after the VR exposure. The changes in intercultural sensitivity between the baseline and right after the VR exposure could be well predicted by the participants' sense of presence and emotional empathy. Moreover, emotional empathy could also strongly predict the retention of the enhancement, while presence was found to be a weaker predictor. These findings, together with the participants' feedback and comments, provided valuable insights and practical implications regarding the future design of VR-enabled intercultural sensitivity training. Meanwhile, the linear regression models could not fully explain the variations in the dependent variables. Thus, given that intercultural sensitivity training could be cultural, geographical, and religious dependent, we suggest that future research may consider other factors (e.g., the sense of embodiment, user experience, and user engagement) and further explore predictive relations on larger and more diverse data sets to further enhance the generalisation of the results.

Acknowledgments

This project is supported by The Hong Kong Polytechnic University (project no.: P0035264) and the Centre for Innovative Applications of Internet and Multimedia Technologies (AIMtech Centre), City University of Hong Kong.

References

[1] M. Akdere, K. Acheson, and Y. Jiang. An examination of the effectiveness of virtual reality technology for intercultural competence development. *International Journal of Intercultural Relations*, 82:109–120, 2021. doi: 10.1016/j.ijintrel.2021.03.009

[2] C. M. Bachen, P. Hernández-Ramos, C. Raphael, and A. Waldron. How do presence, flow, and character identification affect players' empathy and interest in learning from a serious computer game? *Computers in Human Behavior*, 64:77–87, 2016. doi: 10.1016/j.chb.2016.06.043

[3] N. H. Bakker, P. O. Passenier, and P. J. Werkhoven. Effects of head-slaved navigation and the use of teleports on spatial orientation in virtual environments. *Human factors*, 45(1):160–169, 2003. doi: 10.1518/hfes.45.1.160.27234

[4] D. Banakou, P. D. Hanumanthu, and M. Slater. Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in human neuroscience*, 10:601, 2016. doi: 10.3389/fnhum.2016.00601

[5] B. Barbot and J. C. Kaufman. What makes immersive virtual reality the ultimate empathy machine? discerning the underlying mechanisms of change. *Computers in Human Behavior*, 111:106431, 2020. doi: 10.1016/j.chb.2020.106431

[6] M. Bennett. *Basic concepts of intercultural communication: Paradigms, principles, and practices*. Hachette UK, 2013.

[7] M. Bennett. Developmental model of intercultural sensitivity. In *Kim, Y (Ed) International encyclopedia of intercultural communication*. Wiley, 2017.

[8] M. J. Bennett. A developmental approach to training for intercultural sensitivity. *International journal of intercultural relations*, 10(2):179–196, 1986. doi: 10.1016/0147-1767(86)90005-2

[9] P. Bertrand, J. Guegan, L. Robieux, C. A. McCall, and F. Zenasni. Learning empathy through virtual reality: multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. *Frontiers in Robotics and AI*, 5:26, 2018. doi: 10.3389/frobt.2018.00026

[10] D. P. Bhawuk and R. Brislin. The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International journal of intercultural relations*, 16(4):413–436, 1992. doi: 10.1016/0147-1767(92)90031-O

[11] F. P. Brooks. What's real about virtual reality? *IEEE Computer graphics and applications*, 19(6):16–27, 1999. doi: 10.1109/38.799723

[12] Census and Statistics Department, The Government of Hong Kong SAR. Hong Kong 2016 population by-census - thematic report ethnic minorities. `https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=B1120100&scode=459`. Accessed: 2022-04-20.

[13] G.-M. Chen. A review of the concept of intercultural sensitivity. Paper presented at the Biennial Convention of the Pacific and Asian Communication Association (Honolulu, HI, January 1997), 1997.

[14] G.-M. Chen and W. J. Starosta. Intercultural communication competence: A synthesis. *Annals of the International Communication Association*, 19(1):353–383, 1996. doi: 10.1080/23808985.1996.11678935

[15] G.-M. Chen and W. J. Starosta. The development and validation of the intercultural sensitivity scale. *Human Communication*, 3:1–15, 2000.

[16] R. E. Clark. Reconsidering research on learning from media. *Review of educational research*, 53(4):445–459, 1983. doi: 10.3102/00346543053004445

[17] A. J. Coffey, R. Kamhawi, P. Fishwick, and J. Henderson. The efficacy of an immersive 3d virtual versus 2d web environment in intercultural sensitivity acquisition. *Educational Technology Research and Development*, 65(2):455–479, 2017. doi: 10.1007/s11423-017-9510-9

[18] J. Cohen. *Statistical power analysis for the behavioral sciences, 2nd edition*. Erlbaum, 1988.

[19] M. H. Davis et al. A multidimensional approach to individual differences in empathy. 1980.

[20] J. Decety, I. B.-A. Bartal, F. Uzefovsky, and A. Knafo-Noam. Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686):20150077, 2016. doi: 10.1098/rstb.2015.0077

[21] D. Dziedziewicz, A. Gajda, and M. Karwowski. Developing children's intercultural competence and creativity. *Thinking Skills and Creativity*, 13:32–42, 2014. doi: 10.1016/j.tsc.2014.02.006

[22] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner. G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2):175–191, 2007. doi: 10.3758/BF03193146

[23] L. C. Findlay, A. Girardi, and R. J. Coplan. Links between empathy, social behavior, and social understanding in early childhood. *Early Childhood Research Quarterly*, 21(3):347–359, 2006. doi: 10.1016/j.ecresq.2006.07.009

[24] K. M. Gerling, R. L. Mandryk, M. V. Birk, M. Miller, and R. Orji. The effects of embodied persuasive games on player attitudes toward people using wheelchairs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3413–3422, 2014. doi: 10.1145/2556288.2556962

[25] A. Gorini, C. S. Capideville, G. De Leo, F. Mantovani, and G. Riva. The role of immersion and narrative in mediated presence: the virtual hospital experience. *Cyberpsychology, Behavior, and Social Networking*, 14(3):99–105, 2011. doi: 10.1089/cyber.2010.0100

[26] T. Greitemeyer, S. Osswald, and M. Brauer. Playing prosocial video games increases empathy and decreases schadenfreude. *Emotion*, 10(6):796, 2010. doi: 10.1037/a0020194

[27] J. C. Ho and R. Ng. Perspective-taking of non-player characters in prosocial virtual reality games: Effects on closeness, empathy, and game immersion. *Behaviour & Information Technology*, pp. 1–14, 2020. doi: 10.1080/0144929X.2020.1864018

[28] M. L. Hoffman. Interaction of affect and cognition in empathy. *Emotions, cognition, and behavior*, pp. 103–131, 1984.

[29] W. Huang. *Investigating the novelty effect in virtual reality on stem learning*. PhD thesis, Arizona State University, 2020.

[30] S. Jain. Experiential training for enhancing intercultural sensitivity. *Journal of cultural diversity*, 20(1):15, 2013.

[31] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012. doi: 10.1162/PRES_a_00124

[32] M. J. Kors, G. Ferri, E. D. Van Der Spek, C. Ketel, and B. A. Schouten. A breathtaking journey. on the design of an empathy-arousing mixed-reality game. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pp. 91–104, 2016. doi: 10.1145/2967934.2968110

[33] A. Lambert, N. Norouzi, G. Bruder, and G. Welch. A systematic review of ten years of research on human interaction with social robots. *International Journal of Human–Computer Interaction*, 36(19):1804–1817, 2020. doi: 10.1080/10447318.2020.1801172

[34] K. Leung, S. Ang, and M. L. Tan. Intercultural competence. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1):489–519, 2014. doi: 10.1146/annurev-orgpsych-031413-091229

[35] C. Li, H. H. Ip, Y. M. Wong, and W. S. Lam. An empirical study on using virtual reality for enhancing the youth's intercultural sensitivity in hong kong. *Journal of Computer Assisted Learning*, 36(5):625–635, 2020. doi: 10.1111/jcal.12432

[36] M. Lombard and T. Ditton. At the heart of it all: The concept of presence. *Journal of computer-mediated communication*, 3(2):JCMC321, 1997. doi: 10.1111/j.1083-6101.1997.tb00072.x

[37] M. Luck and R. Aylett. Applying artificial intelligence to virtual reality: Intelligent virtual environments. *Applied artificial intelligence*, 14(1):3–32, 2000. doi: 10.1080/088395100117142

[38] D. Lussier. Theoretical bases of a conceptual framework with reference to intercultural communicative competence. *Journal of Applied Linguistics*, 4(3), 2007. doi: 10.1558/japl.v4i3.309

[39] G. Makransky, T. S. Terkildsen, and R. E. Mayer. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60:225–236, 2019. doi: 10.1016/j.learninstruc.2017.12.007

[40] A. J. Martingano, F. Hererra, and S. Konrath. Virtual reality improves emotional but not cognitive empathy: A meta-analysis. *Technology, Mind, and Behavior*, 2(1), 2021. doi: 10.1037/tmb0000034

[41] A. Mehrabian and N. Epstein. A measure of emotional empathy. *Journal of personality*, 1972.

[42] Z. Merchant, E. T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, and T. J. Davis. Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: A meta-analysis. *Computers & Education*, 70:29–40, 2014. doi: 10.1016/j.compedu.2013.07.033

[43] N. Norouzi, K. Kim, J. Hochreiter, M. Lee, S. Daher, G. Bruder, and G. Welch. A systematic survey of 15 years of user studies published in the intelligent virtual agents conference. In *Proceedings of the 18th international conference on intelligent virtual agents*, pp. 17–22, 2018. doi: 10.1145/3267851.3267901

[44] T. C. Peck, S. Seinfeld, S. M. Aglioti, and M. Slater. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition*, 22(3):779–787, 2013. doi: 10.1016/j.concog.2013.04.016

[45] Ş. Penbek, D. Yurdakul Şahin, and A. G. Cerit. Intercultural communication competence: A study about the intercultural sensitivity of university students based on their education and international experiences. *International Journal of Logistics Systems and Management*, 11(2):232–252, 2012. doi: 10.1504/ijlsm.2012.045425

[46] T. Piumsomboon, Y. Lee, G. A. Lee, A. Dey, and M. Billinghurst. Empathic mixed reality: Sharing what you feel and interacting with what you see. In *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*, pp. 38–41. IEEE, 2017. doi: 10.1109/ISUVR.2017.20

[47] R. S. Rosenberg, S. L. Baughman, and J. N. Bailenson. Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PloS one*, 8(1):e55003, 2013. doi: 10.1371/journal.pone.0055003

[48] M. Şahin. An investigation into the efficiency of empathy training program on preventing bullying in primary schools. *Children and Youth Services Review*, 34(7):1325–1330, 2012. doi: 10.1016/j.childyouth.2012.03.013

[49] A. Q. Sarwari and M. N. Abdul Wahab. Study of the relationship between intercultural sensitivity and intercultural communication competence among international postgraduate students: A case study at university malaysia pahang. *Cogent Social Sciences*, 3(1):1310479, 2017. doi: 10.1080/23311886.2017.1310479

[50] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. doi: 10.1162/105474601300343603

[51] M. J. Schuemie, P. Van Der Straaten, M. Krijn, and C. A. Van Der Mast.

Research on presence in virtual reality: A survey. *CyberPsychology & Behavior*, 4(2):183–201, 2001. doi: 10.1089/109493101300117884

[52] G. Serafin and S. Serafin. Sound design to enhance presence in photorealistic virtual reality. In *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display*, 2004.

[53] D. Shin. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78:64–73, 2018. doi: 10.1016/j.chb.2017.09.012

[54] N. R. Silton, E. Fontanilla, M. Femia, and K. Rouse. Employing disability simulations and virtual reality technology to foster cognitive and affective empathy towards individuals with disabilities. In *Scientific Concepts Behind Happiness, Kindness, and Empathy in Contemporary Society*, pp. 191–207. IGI Global, 2019. doi: 10.4018/978-1-5225-5918-4.ch010

[55] R. Skarbez, F. P. Brooks, Jr, and M. C. Whitton. A survey of presence and related concepts. *ACM Computing Surveys (CSUR)*, 50(6):1–39, 2017. doi: 10.1145/3134301

[56] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557, 2009. doi: 10.1098/rstb.2009.0138

[57] M. Slater. Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109(3):431–433, 2018. doi: 10.1111/bjop.12305

[58] M. Slater and M. Usoh. Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, 2(3):221–233, 1993. doi: 10.1162/pres.1993.2.3.221

[59] M. Slater, M. Usoh, and A. Steed. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 3(2):130–144, 1994. doi: 10.1162/pres.1994.3.2.130

[60] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, 1992. doi: 10.1111/j.1460-2466.1992.tb00812.x

[61] D. A. Straffon. Assessing the intercultural sensitivity of high school students attending an international school. *International journal of intercultural relations*, 27(4):487–501, 2003. doi: 10.1016/S0147-1767(03)00035-X

[62] R. M. Tamim, R. M. Bernard, E. Borokhovski, P. C. Abrami, and R. F. Schmid. What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational research*, 81(1):4–28, 2011. doi: 10.3102/0034654310393361

[63] C. Tarchi, A. Surian, and C. Daiute. Assessing study abroad students' intercultural sensitivity with narratives. *European journal of psychology of education*, 34(4):873–894, 2019. doi: 10.1007/s10212-019-00417-9

[64] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence*, 9(5):497–503, 2000. doi: 10.1162/105474600566989

[65] R. Venkatakrishnan, R. Venkatakrishnan, R. G. Anaraky, M. Volonte, B. Knijnenburg, and S. V. Babu. A structural equation modeling approach to understand the relationship between control, cybersickness and presence in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 682–691. IEEE, 2020. doi: 10.1109/VR46266.2020.00091

[66] S. Ventura, L. Badenes-Ribera, R. Herrero, A. Cebolla, L. Galiana, and R. Baños. Virtual reality as a medium to elicit empathy: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 23(10):667–676, 2020. doi: 10.1089/cyber.2019.0681

[67] V. T. Visch, E. S. Tan, and D. Molenaar. The emotional and cognitive effect of immersion in film viewing. *Cognition and Emotion*, 24(8):1439–1445, 2010. doi: 10.1080/02699930903498186

[68] P. Vorderer, S. Knobloch, and H. Schramm. Does entertainment suffer from interactivity? the impact of watching an interactive tv movie on viewers' experience of entertainment. *Media Psychology*, 3(4):343–363, 2001. doi: 10.1207/S1532785XMEP0304_03

[69] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998. doi: 10.1162/105474698565686

[70] X. Zhang and M. Zhou. Interventions to promote learners' intercultural competence: A meta-analysis. *International journal of intercultural relations*, 71:31–47, 2019. doi: 10.1016/j.ijintrel.2019.04.006