Designing and Assessing a Virtual Reality Simulation to Build Resilience to Street Harassment

Crescent Jicol
c.jicol@bath.ac.uk
University of Bath
Bath, UK

Julia Feltham
jf809@bath.ac.uk
University of Bath
Bath, UK

Jinha Yoon
jhy33@bath.ac.uk
University of Bath
Bath, UK

Michael J Proulx
m.j.proulx@bath.ac.uk
University of Bath
Bath, UK

Eamonn O'Neill
e.oneill@bath.ac.uk
University of Bath
Bath, UK

Christof Lutteroth
c.lutteroth@bath.ac.uk
University of Bath
Bath, UK

Figure 1: We co-designed a VR simulation allowing users to experience a typical street harassment situation in an immersive and interactive environment (a). While waiting in a street at night (b), the user is approached by a male harasser (c). We investigated VR design choices and their effects on realism and effectiveness, as well as the surrounding ethical issues (d).

ABSTRACT

Street harassment is a widespread problem that can constrain people's freedom to enjoy public spaces safely, along with many other negative psychological impacts. However, very little research has looked at how immersive technology can help in addressing it. We conducted three studies to investigate the design decisions, ethical issues and efficacy of an immersive simulation of street harassment: an online design study (n=20), an interview study with experts working in the area (n=9), and a comparative lab study investigating design, ethics and efficacy (n=44). Our results deepen understanding of the design decisions that contribute to a realistic psychological experience, such as the effects of screen-based video *vs* passive VR *vs* interactive VR. They also highlight important ethical issues such as traumatisation and potential for victim blaming, and how they can be approached in an ethical manner. Finally, they provide insights into efficacy in terms of perceived usefulness, competence and empathy.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3502129

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

street harassment, virtual reality, design, ethics

ACM Reference Format:
Crescent Jicol, Julia Feltham, Jinha Yoon, Michael J Proulx, Eamonn O'Neill, and Christof Lutteroth. 2022. Designing and Assessing a Virtual Reality Simulation to Build Resilience to Street Harassment. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3491102.3502129

1 INTRODUCTION

Street harassment is a serious global problem that affects most women. The vast majority of women have experienced it before the age of 17 (e.g. 90% of women in the UK) [5, 43, 48]. This issue includes not only verbal harassment but also even more serious instances, such as stalking and sexual assault: 71% of women globally reported being followed and 50% reported being fondled or groped. In a US study, two-thirds of the harassed women were concerned that the incident would escalate into something worse, with 41% having experienced physically aggressive harassment [43]. Many women experience street harassment frequently, with 43% of participants in a US study [16] reporting "sexist remarks or behaviours" and 14.8% reporting "unwanted touching, stroking, or hugging" at

least once a month. Men also experience street harassment [37, 43], especially if they do not fit a binary gender stereotype [21, 62]. Gardner describes street harassment as "a continuum of possible events, beginning when customary civility among strangers is abrogated and ending with the transition to violent crime" [32].

Street harassment has severe negative impacts. The fear, anger and frustration generated by its prevalence affects many people's everyday lives and constrains their freedom to enjoy public spaces safely. The fear of a street harassment situation escalating into an even more serious situation is a main reason for its strong psychological effect. According to Bowman, fear of rape is one of the two most common themes that appear when women discuss their harassment experiences, with the other being intrusion on privacy [10]. Many women report regularly taking steps to avoid street harassment (e.g. 61% in the US) [99], such as changing the routes to their destinations (e.g. 88% in Italy) or being unwilling to go out at night at all (e.g. 80% in India) [48]. Street harassment has been found not only to restrict people's geographic mobility but also to have wider psychological impacts such as feelings of loss of control and reduced self-esteem [19]. It has been linked to generalised anxiety [16], such as "feeling unable to relax and fearing the worst happening as well as physiological responses such as feeling unsteady, hands trembling". Perceptions of safety in isolated public settings could thus act as an important mediator in the relationship between street harassment and anxiety. Self-objectification where people "internalise an observer's perspective as a primary view of their physical selves" is also strongly linked to street harassment [29]. Self-objectification can lead to shame, reduce motivation and academic achievement, and lead to mental health problems including anxiety, depression, sexual dysfunction and eating disorders [18, 43].

Some technologies have been proposed to help address street harassment and its negative impacts. Personal safety mobile apps offer alarms to draw the attention of passers-by and call for help, dedicated maps to avoid harassment hot-spots, and forums for sharing information and experiences [41, 42, 98]. While such apps can be useful, they do not address one of the main challenges faced by harassment victims: psychological stress. Most people find it difficult to make decisions and act in such stressful situations [23, 81]. In extreme cases such as rape or attempted rape, 81.6% of survivors have reported experiencing immobility, i.e. "freezing" and being unable to think clearly or act [63]. Immersive applications have been proposed [1–3, 39, 40, 85] to help people prepare for the psychological impact of experiencing harassment. Based on the long-established practices of stress inoculation training [58] and role play [9, 89], these applications allow people to experience minor harassment stressors to foster psychological preparedness and promote resilience, e.g. by developing awareness and coping strategies. However, there is little published research on simulating harassment with immersive technology, especially for street harassment. We address this gap by investigating three challenges in the design, use and evaluation of such immersive simulations:

Challenge 1: For harassment simulations to be effective they need to be realistic [40]. This is a challenging requirement in staged *in vivo* role plays [8], and immersive technology holds promise in addressing it: there is some evidence that immersive *in virtuo* simulations of harassment can deliver sufficient realism in a consistent and safe manner [39, 40, 51, 52, 64, 91]. However, while some of the effects of immersive harassment simulations have been described before, there is no published research investigating the design decisions that make such simulations realistic. We address this gap by eliciting design feedback from users and experts, highlighting important design decisions and measuring their effects in a comparative user study.

Challenge 2: Ethical issues arise when simulating street harassment. One such issue is the risk of simulations triggering reactions due to past trauma. Another, more subtle issue is that immersive harassment 'training' can sometimes be perceived as 'victim blaming', i.e. as shifting the onus of dealing with the problem of harassment to the victim. Ethical issues of simulating street harassment with immersive technology are neglected in previous literature, with brief mentions of ethical approvals and exclusion criteria being the only relevant references. We present the first analysis of ethical issues in simulating street harassment based on interviews with experts and users, and provide design recommendations to address these issues in research and practice.

Challenge 3: Evaluating the effectiveness of simulations of harassment in preparing users for real harassment is difficult. On the one hand, it is uncertain how far lab measurements of success such as observed behaviours translate into the real world. On the other hand, exposing participants to real harassment, or even the perception of real harassment, to evaluate preparedness is dangerous and unethical. Published studies on immersive simulations of harassment largely ignore the question of real world effectiveness [39, 40, 51] or focus on measures of empathy towards victims as indicators of success [36, 64, 91]. In a notable exception, Loucks et al. [52] evaluated the efficacy of VR exposure therapy to treat sexual harassment trauma in the military using long term mental health outcomes. However, no published study has considered real world effectiveness of street harassment simulations. We aim to address this by using quantitative and qualitative measures of effectiveness with both users and experts, with a focus on the effectiveness of a simulation in eliciting realistic emotions, and consider their implications for design.

We investigate these challenges through the following research questions:

- **RQ1** How can we design a realistic simulation of street harassment?
- **RQ2** What are the ethical issues surrounding simulations of street harassment?
- **RQ3** How effective is a simulation of street harassment in eliciting realistic emotions?

We tackle these research questions by co-designing an immersive simulation of street harassment (Figure 1) and evaluating it in three user studies. Study 1 is an online design study, Study 2 is an interview study with experts working in the area of harassment, and Study 3 is a comparative user study. While Study 1 addresses mainly RQ1, Studies 2 and 3 address all three research questions. In summary, we make the following contributions:

(1) An evidence-based immersive street harassment simulation.
(2) A critical analysis of the ethical issues surrounding immersive simulations of street harassment, with design recommendations to address them.

(3) Empirical insights into the effectiveness of such immersive simulations in eliciting realistic emotions, based on a mixture of quantitative and qualitative methods.

2 RELATED WORK

Stress Inoculation Training (SIT) is used to "prepare individuals for stressful situations (such as combat or medical emergencies), diminishing the potential for a negative psychological reaction like PTSD" [71, 96]. SIT "aims to build stress tolerance through exposure" [65] and, as in exposure therapy, this is achieved through the "gradual, controlled, and repeated exposure to a stressor". This exposure aims to prevent a panicked or 'fight or flight' response to a real situation through desensitisation [96], increasing the user's resilience and control. Unlike exposure therapy, which focuses on personal phobias, SIT focuses on environments which are stressful to most people. A study in the military [97] found that VR SIT can reduce a participant's emotional responses to negative stimuli even three months after experiencing it. One way to administer SIT is through role-play. Role-playing exercises have been used in behavioural and cognitive-behavioural therapies for a long time and have proved successful at helping individuals develop strategies for coping with social situations that are complex [35, 73, 76, 83]. They have also been proposed in teaching effective strategies to deal with sexual harassment [39, 72], however, there has been little research in the development and evaluation of role-playing for this purpose.

VR Design of Realistic Stressors. VR can be used to enhance the effect of SIT, by providing the user with a sense of presence and evoking reactions and emotions similar to a real experience [78]. A review of six studies [96] concluded that VR-enhanced SIT is "more effective than real world training, in terms of time expenditure as well as helping participants adapt to stressful stimuli and performing efficiently". In the context of exposure therapy, VR has been found to have beneficial effects close to those of *in vivo* exposure while offering practical advantages [95]. More specifically, Juriles et al. [39, 40] showed that VR could enhance the realism of role-plays designed to help women resist sexual assaults. Based on self-reports of negative affect and perception of realism, direct observation of participants' verbal displays of negative affect, and heart rate measurements, participants perceived verbal threats and sexual advances to be more realistic in VR than in *in vivo*. However, similar to other SIT studies, the researchers did not explore why VR was apparently so effective and did not discuss VR design aspects. Agency in particular is an aspect of VR design that is known to have a substantial effect on presence [20], especially for VR stimuli inducing negative emotions [38]. Affording users agency has been shown to enhance physical presence, or the impression of being actually surrounded by landmarks in a virtual environment (VE) [44]. A sense of agency can also enhance the impression that one is present in a VE, referred to as self-presence [15], and the feeling of being in the presence of other living beings in the VE [69]. This raises the question of how far agency can heighten presence and affect emotions in SIT for street harassment, by allowing the user to act in the virtual environment.

Immersive Harassment Simulations. There are some commercial role-play based VR training tools for sexual harassment. Vantage Point [3] and Regatta [2] immerse users in 360° videos of sexual harassment scenarios, performed by actors in a variety of workplace settings. However, these tools focus on training the user in the role of a bystander rather than the victim, and they lack scientific validation. Some research has shown that immersive harassment simulations can increase empathy in men for female victims, which is important because of the link between lack of empathy and antisocial behaviour such as harassment [56]. Ventura et al. [91] compared a 360° video with a traditional non-immersive narrative, each placing male participants in the role of a female victim. Both conditions were effective in increasing empathy for female victims and decreasing violent attitudes towards women, with the 360° video performing better than the non-immersive narrative. Seinfeld et al. [79] used VR to embody male participants in a female body, who then received verbal abuse from a male avatar. The male participants subsequently showed increased empathy towards female facial stimuli depicting fear. Neyret et al. [64] used immersive VR to let male participants embody either an observer or a female victim in a sexual harassment scenario. The men who experienced the scenario as the victim showed more empathy and less aggressive behaviour in a follow-up virtual Milgram experiment [60] a week later.

While related work indicates that VR holds promise for applying SIT to street harassment, it does not provide answers to our research questions. Immersive SIT training tools focus mainly on workplace harassment [2, 3], and studies involving immersive harassment simulations focus mainly on their effects on male participants, notably empathy [64, 91]. While the importance and complexity of design factors such as presence [78] and agency [38] have been recognised in the wider VR literature, the question of how realistic street harassment simulations can be designed (RQ1) has hardly been addressed at all.

Regarding the ethical issues of simulating street harassment (RQ2), while there are suggestions that VR simulations can have adverse effects [64], ethical issues have generally not been explored. What previous work there is suggests that the risk of triggering strong emotional reactions is low, with a sexual assault VR exposure therapy study [51] and two sexual harassment VR role-play studies [39, 40] not reporting any severe emotional responses or dropouts due to participants feeling too uneasy. However, while some ethical issues in simulated abuse have been discussed in the context of social human-robot interaction experiments [67], there exists no research showing how safety and other ethical considerations should be considered in a VR design.

Perhaps the most challenging research question is on the effectiveness of VR in psychologically preparing people by eliciting realistic emotions felt during street harassment situations (RQ3). Previous work has addressed this mainly from the perspective of men, using VR to elicit empathy for female victims [64, 91]. This is important but does not consider other perspectives, and does not contribute directly to our understanding of the appropriateness and efficacy of interventions intended to prepare people psychologically for being harassed. In the following sections, we describe three studies which together aim to address each of these gaps in the literature.

3 STUDY 1: ONLINE DESIGN STUDY

Our first study focused on co-designing a VR street harassment simulation with experts and end users, and eliciting design feedback through videos and an online questionnaire. It addressed RQ1 by providing evidence to help designers overcome the "fundamental source of invalidity" in *in vivo* role-plays, which is that people often do not respond in a "natural fashion" due to the failure of the interaction to feel vivid and realistic [8]. The co-design process was guided by previous related work and involved end users and a harassment expert.

3.1 VR Harassment Scenario

Environment. For the simulation to be ecologically valid, it was important to base the virtual role-play on common characteristics of real street harassment situations. A 2018 report [34] on girls' safety across five cities found street harassment was most prevalent in urban environments near transport hubs such as train stations and bus stops, with a peak in harassment activity late in the evening and at night. Another study [26] looking at the most common words used in real harassment stories supports this report, finding that 'street', 'bus' and 'train' were all in the top 15, along with 'night', which was more common than 'day'. 'Waiting' was also a word frequently found, suggesting that it is common for a person to experience harassment when waiting for a bus in a public space. Conversations with potential end users and a harassment expert confirmed that waiting for public transport at night was a common situation particularly laden with fear of harassment. As a result, the environment (Figure 1) was designed to present an urban street where the user is waiting at a bus stop at night. It was built with Unity using off-the-shelf assets from the Unity Store.

Harasser. Being harassed by one man was the most common experience of both men and women [48]. Therefore, a single male avatar was chosen to play the role of the harasser. The harasser was designed to be wearing dark glasses as these help prevent an uncanny valley effect occurring due to an avatar's unrealistic eye gaze [33]. The harasser was animated using a library of pre-recorded movements which were used to match the script. The harasser was programmed to turn according to where the user was facing to increase realism. Also to increase realism, the harasser's voice was played by a professional actor. The voice script (see Appendix 1) included selected phrases from a list of verbal harassment lines by Livingston et al. [48] in order to increase ecological validity. The harasser's voice was recorded to match the movements of the avatar, and the avatar's jaw and lips were animated to move in sync with the voice.

Role Play. The simulation was designed to involve verbal harassment but not physical harassment. Verbal harassment is more common than physical harassment [48], and simulating physical harassment would likely break immersion due to extreme levels of emotions felt and cause serious ethical concerns. It was important to ensure that the actions of the virtual harasser appeared as harassment instead of, for example, "consensual flirting, polite hellos, and respectful small talk" [48]. To design the content of the role-play accordingly, common characteristics of verbal street harassment situations were researched. Other studies involving VR sexual harassment role-play to induce realistic emotions [39, 40] split their role-play into phases, with each successive phase increasing in intensity. Drawing on this previous work, we systematically increased the intensity in our scenario, by (a) making the tone of the harasser's voice get louder, (b) moving the harasser closer to the victim, and (c) making the harasser's movements more vigorous and pronounced (e.g. pacing around, raising hands) as the experience progressed.

The simulation proceeds as follows: (1) the user is given some time to adapt to their new environment, with only background activity such as other city dwellers chatting; (2) the user is catcalled by the approaching harasser; (3) the harasser walks towards the user; (4) he stares at the user, catcalls, asks invasive questions and tries to get their attention; (5) he steps closer and persistently asks them to go on a date with him; (6) he starts to become upset that they are not willing to go with him; (7) he becomes angry at the rejection; (8) he leaves the scene. The entire simulation was designed to last 5 minutes in order for the user to develop a sense of presence and potentially experience the full onset of a stress response, while limiting the possibilities of overexposure or boredom [101].

Interactivity. Affording agency implied allowing the user to exert some control over the threatening stimulus. The role-play dialogue was designed so there were pauses after the harasser said phrases, giving the user an opportunity to respond if they wished. Previous work [39, 40] suggested that interaction with harassers is usually marked by raising one's voice. If the user raised their voice above 70 decibels, which is above the volume of a typical conversation, then this was considered an interaction. The user could also interact with the harasser by putting their hands out in a gesture to stop. Every time the user interacted with the harasser, either by raising their voice or raising their hands, the harasser was programmed either to take a step back or to make a comment such as "being cocky are we?". This variation in reactions was implemented to increase realism as humans may not react the same way repeatedly. Each interaction was pilot tested with five participants, resulting in the 70 decibel threshold for the voice interaction. For the hand interaction, it was found that the harasser needed to be within 2 metres of the victim for raising one's hands to feel like a natural defensive reaction. This too was important for maintaining a sense of realism during the interaction.

3.2 Evaluation

Methodology. We recruited 20 participants (3 males, 17 females), aged 21-42 ($M = 23.65, SD = 5.16$) through social media to provide feedback on our first design prototype in an online questionnaire study. Because of COVID-19 restrictions it was conducted online, with participants watching a 2 minute video showing the main parts of the simulation and its interactive features from a first-person perspective on their own screens. Some participants used mobile devices and some used desktop or laptop computers. Given these limitations, our main focus was on assessing individual design elements and gathering formative design feedback rather than assessing the complete immersive experience.

After viewing the video, participants rated their perceived presence in the simulation scenario on a 7 point scale (7=best) using the single item presence subscale and two additional items for realism and physical presence from the iGroup Presence questionnaire

(IPQ) [77]. This was followed by an open question about the realism of the environment, three 7 point rating items about the harasser (*'How real did the harasser seem to you?'*, *'The harasser gave me feelings of eeriness.'*, *'The harasser appeared creepy.'*), and an open question about the harasser. Finally, 5 point rating questions asked how far presence was increased by specific interactive features: (a) programming the harasser always to turn to face the user, (b) allowing the user to see their own hands, and having the harasser step back or respond verbally when the user raises (c) their voice or (d) their hands. The rating items were all based on the IPQ.

Measuring presence for non-immersive videos is a fairly common practice, even for short videos [24, 59]. Presence scores are generally lower for 2D videos compared to more immersive formats, but still considered meaningful and comparable [92]. The IPQ has been widely used for that purpose [30, 47, 50], and a database of research using the IPQ indicates that the majority of studies have measured presence with content presented on a normal 2D screen [90].

Results. Participants felt physical presence in the VE ($M = 5.5, SD = 1.32$), and also perceived fairly high degrees of self-presence ($M = 5.15, SD = 1.69$) and realism ($M = 4.25, SD = 1.11$) (*"Felt completely immersed which is amazing given that I wasn't even watching it through the headset, just on my phone."*). The environment was perceived as dark, enclosed and unsafe by participants (*"dingy back alley"*), adequately setting the scene for an emotionally stressful experience (*"fitted the situation and 'the vibe' that was intended perfectly"*). The most common suggestion for improving the realism of the environment was to add some more people in the background or walking past. Another common suggestion was to add more ambient noise such as traffic and sirens. Some suggested adding a brief narrative to aid immersion at the beginning (*"Imagine you are walking home after meeting friends ..."*). Participants felt that the harasser appeared real ($M = 4.8, SD = 0.89$) and 'creepy' ($M = 6.30, SD = 0.73$), giving them feelings of eeriness ($M = 6.15, SD = 1.04$). Several participants commented positively on the convincing voice acting (*"very good voice acting"*, *"very similar to things I've heard in real life"*) and the overall realistic behaviour of the avatar (*"the avatar and his responses are really good, makes it feel much more like you're 'in the experience' rather than 'watching the experience' to me"*), while some suggested that some of the movements could be more natural. Several participants suggested making the avatar taller (*"The harasser being of larger build would be more threatening."*). All the interactive features were perceived as effective by respondents in terms of making users feel more present in the virtual environment: harasser facing the user ($M = 4.55, SD = 0.76$), users seeing their own hands ($M = 4.15, SD = 1.14$), harasser reacting to voice ($M = 4.35, SD = 0.93$), and harasser reacting to raising hands ($M = 4.3, SD = 0.73$).

Discussion. The results from Study 1 suggest that the design could be effective in recreating a realistic harassment experience (RQ1). Drawing on these results, we refined our design based on the participants' suggestions. The increased realism and ecological validity of the tool could be instrumental in effective training but also pose risks to the mental well being of users, hence, Study 2 was conducted to better understand ethical concerns so they can be addressed before the tool is deployed.

4 STUDY 2: EXPERT INTERVIEW STUDY

To understand the ethical issues around (RQ2) and the effectiveness of (RQ3) immersive VR role-play of street harassment, we interviewed experts in the field. Study 2 involved conducting 30 minute online video interviews with participants who worked in a professional role providing support to street/sexual harassment victims. Participants were sent a demo video of the VR tool prior to the interview that included the main street harassment role-play and demos of the interactive features. Then, questions were asked to evaluate the simulation, with a particular emphasis on the ethical issues it may pose if used to psychologically prepare people for real street harassment.

Participants. Interviews were conducted with nine experts with diverse experience including charity work with victims of street harassment, training people in bystander intervention and instructing empowerment self-defence (ESD), including a co-founder of an ESD centre and a self-defence coordinator who has taught self defence for decades. Recruitment was conducted by advertising via the websites and social media of sexual harassment charities/organisations and empowerment self-defence organisations.

4.1 Results

The interview data were analysed using a thematic analysis based on the six-step process described by Maguire et al. [54].

4.1.1 Realism. All interviewees felt that the virtual role-play realistically simulated a street harassment scenario, particularly in terms of what the harasser said and also the way the harasser escalated and turned on the victim. However, it was noted that this was only one of many possible scenarios (*"generally I thought it reflected one possibility"*). A few participants also mentioned they felt emotions just watching the simulation on a screen (*"Even not doing it through VR I felt quite uncomfortable and felt a bit nervous"*, *"The adrenaline one feels watching even that situation, it definitely relates to what I feel on the street sometimes"*). There were mixed opinions on whether a person who has never experienced harassment would feel such emotions when experiencing the simulation.

4.1.2 Ethics. Several participants suggested that the tool may trigger emotions associated with street harassment that users had experienced (*"I think what the scenario has the ability to do is to trigger emotions that people have already experienced"*). Because of this, some participants agreed that there is a potential risk to some users, especially those with PTSD related to harassment experiences (*"it would not re-traumatise them [users with PTSD], I don't think, but it could definitely bring up some overwhelming emotions"*). All participants felt that this risk could be mitigated by adding some features to the tool or by using the tool in the presence of professionals. Suggested features included guidelines on how to use the tool safely, a clear description of what the tool entails, a video/activity for users to do prior to the simulation to assess if it is safe for them to use, a way easily to exit the simulation at any time, and a way for the user to receive support if needed after using the tool (e.g. provide numbers of help lines to call).

Some participants voiced a concern that the simulation may be perceived by some people as 'victim blaming', i.e. as shifting the responsibility for harassment to the victims by 'training' them.

The majority of participants did not view the simulation as victim blaming but explained how they could see others perceiving it in that way, especially if it was portrayed as solely for women (*"I believe there will be people who see it as victim blaming, simply because you are trying to provide tools to women"*). All participants that had experience in ESD instructing mentioned that ESD is viewed in that way by some people, as it is intended for training women. Participants suggested that to avoid the simulation being perceived as victim blaming, it should be made explicit that the harasser is to blame for the situation (*"In all the SD work, we make it very clear it is the perpetrator who is responsible"*), by including text saying it is not their responsibility to protect themselves and that the blame is always solely on the harasser (*"could even have a message at the end to clarify, you are not to blame for anything that occurs, your reactions are completely natural, etc. The blame solely is with the harassers"*). Participants also suggested that the tool could be interpreted less as victim blaming if it were used in more ways, such as for bystander training [31, 64] or as an awareness raising tool for both potential victims and potential harassers.

Participants emphasised that there is a wide range of possible harassment scenarios, and that different scenarios require different behaviours when dealing with them. One participant was concerned that a simulation would train users in particular behaviours that may be inappropriate in some cases (*"some situations can be dangerous"*). She advised that such a simulation should not prescribe or explicitly encourage specific behaviours as general responses.

4.1.3 Effectiveness. Participants agreed that the simulation could be used to help prepare people for real harassment scenarios (*"a strong tool"*, *"could make training more practical and real"*). However, participants emphasised that it reflected only one scenario and to successfully prepare people for real scenarios, a wider variety of simulated scenarios should be supported. Some participants suggested the ability to adapt the scenario to the demographic of the user, as this can affect the type of street harassment received in the real world.

4.2 Discussion

The results suggest that the simulation presents at least one common street harassment scenario realistically (RQ1). While there were several ethical concerns such as the potential for training being interpreted as victim blaming (RQ2), participants felt that they could be addressed adequately with adjustments to the simulation design and the procedure surrounding its use. Participants agreed that the simulation can be effective in eliciting realistic emotions (RQ3) but that this comes with potential risks. Ultimately, they agreed that the simulation could be used safely, ethically and effectively to prepare people psychologically for street harassment. We integrated many of the participants' suggestions, e.g. regarding ethics, iterating the design before testing the simulation with users in Study 3.

5 STUDY 3: COMPARATIVE USER STUDY

In order to further investigate our research questions, we conducted a controlled experiment with potential end users using a mixed methods approach including both quantitative and qualitative measures. This study tested several versions of the tool to provide insights into the effects of VR design decisions, looking at agency, presence, emotion, ethics and effectiveness in preparing people for street harassment. The study was approved by the Department of Psychology Ethics Committee at the University of Bath (ethics code: 21-192).

5.1 Methodology

Some previous work has suggested that VR can offer advantages over 2D screens in some training simulations [82]. We designed the study to shed light on any such advantages for a simulation of street harassment. Furthermore, it has been suggested that affording users agency in VR could increase their sense of presence [20, 80, 100], which in turn may increase the intensity of elicited emotions [7, 70]. We investigated this possibility by considering two versions of the VR harassment simulation: one with ('active') and one without ('passive') interactive features. We expected that VR would elicit higher levels of presence and emotions compared to a screen, and that affording agency in VR would further increase these effects. Lastly, previous studies on immersive *in virtuo* simulations of harassment have not considered gender differences [39, 40, 51, 52, 64, 91]. We tested both male and female participants to explore some of the effects of gender when simulating street harassment.

We used a within participants design with the order of the conditions counterbalanced. In the screen condition (S), participants watched a pre-recorded video of the simulation on a screen. In the passive VR condition (PVR), participants experienced the simulation in VR without the interactive features. In the active VR condition (AVR), participants experienced the simulation in VR with the interactive features, having been informed that they could verbally respond to the harasser and could put their hands out in front of them when the harasser gets close if they felt comfortable doing so.

5.1.1 Outcome Measures. Presence was measured after every condition using the well validated Multimodal Presence Scale (MPS) [55], which has subscales for physical presence, social presence and self presence. The MPS consists of 15 items rated on 5 point Likert scales, yielding overall scores between 1 (least presence) and 5 (most presence). Agency was measured using the questionnaire used by Jicol et al. [38], based on the well validated "User Experience in Immersive Virtual Environments" questionnaire [88]. It comprises three items rated on a 10 point Likert scale from 1 (least agency) to 10 (most agency).

Emotional response was measured using multiple measures. The negative affect subscale of the well validated and widely used Positive and Negative Affect Schedule (PANAS) [94] was used to measure an emotional baseline at the start, as well as after every condition. The subscale consisted of five emotionally negative items (afraid, upset, nervous, ashamed, hostile), rated on a 5 point scale, yielding an overall score from 1 (most positive) to 5 (most negative). Additionally, as harassment is known to cause anxiety [16, 29, 43], the 6 item short form of the State-Trait Anxiety Inventory (STAI) [84] was used to measure an emotional baseline at the start and after every condition. It was scored using a 5 point Likert scale from 1 (least anxiety) to 5 (most anxiety). Heart rate (HR), which is a common indicator of emotional stress [87] and arousal [6, 22], was measured during every condition using a Polar 10 chest-strap

HR monitor and aggregated by averaging over the experience. Skin conductance, also known as electrodermal activity or galvanic skin response, is another common physiological indicator of emotional stress [61]. We measured it in every condition using a Shimmer Consensys GSR Development Kit with electrodes on the middle and index fingers of the non-dominant hand, and averaged over the duration of the experience.

We measured a participant's empathy for harassment victims at baseline (before starting the experimental conditions), after every condition and one day after the experiment, using the empathy scale from Ventura et al. [91]. The questionnaire comprises five items rated on a 5 point scale that assess a participant's ability to take the perspective of a harassment victim, yielding a score from 1 (least empathy) to 5 (most empathy). We also used the Value/Usefulness, Perceived Competence and Pressure/Tension subscales of the well validated and widely used Intrinsic Motivation Inventory (IMI) [17, 49] after every condition, to measure how much participants felt the simulation would be useful in preparing them for real street harassment scenarios (Value/Usefulness), how competent they felt in handling the harassment situation (Perceived Competence), and how much pressure participants felt during the experience (Pressure/Tension). IMI items were rated on a 7 point scale, with subscale scores from 1 (least) to 7 (most).

At the end of the study, we administered a mixture of study-specific ranking questions, rating questions and open ended questions about the realism, effectiveness and ethics of the simulation. Participants ranked the conditions from most to least realistic and explained their ranking and overall opinion of realism. Four open-ended questions (*"What use do you think this VR harassment simulation could have?"*, *"What use, if any, could this tool have in helping to reduce street harassment?"*) and 7 point Likert scale items (*"If I were to regularly experience the VR harassment situation, it would make me feel more psychologically prepared for a future real-life situation"*, *"If I were to regularly experience the VR harassment situation, it would make me feel more afraid of a future real-life situation"*) assessed the effectiveness of the simulation. Lastly, two open ended questions and a 7 point Likert scale item (*"I think VR harassment simulations like this are ethical"*) were aimed at understanding participants' views on ethical issues. Due to the sensitive nature of street harassment, a written rather than oral response was used for all questions to mitigate participants' reluctance in sharing their personal experiences. The full post-questionnaire can be found in Appendix 2 in the supplementary material.

5.1.2 Hypotheses. Based on related work and pilot experiments, we posed *a priori* hypotheses that our harassment simulations become more effective with increasing immersiveness and realism from S to PVR and further to AVR. Each hypothesis H1-H10 has two sub-hypotheses, e.g. H1P and H1A, addressing the expected gains of PVR *vs* S and AVR *vs* PVR respectively. The first three hypotheses describe gains in presence:

- **H1** Physical presence is higher in PVR *vs* S (**H1P**) and higher in AVR *vs* PVR (**H1A**).
- **H2** Social presence is higher in PVR *vs* S (**H2P**) and higher in AVR *vs* PVR (**H2A**).
- **H3** Self presence is higher in PVR *vs* S (**H3P**) and higher in AVR *vs* PVR (**H3A**).

The next five hypotheses describe an increasing capacity to induce emotional stress, as measured by subjective scales (H4-H6) and physiological measures (H7 and H8):

- **H4** PANAS scores are higher in PVR *vs* S (**H4P**) and higher in AVR *vs* PVR (**H4A**).
- **H5** STAI scores are higher in PVR *vs* S (**H5P**) and higher in AVR *vs* PVR (**H5A**).
- **H6** IMI Pressure/Tension scores are higher in PVR *vs* S (**H6P**) and higher in AVR *vs* PVR (**H6A**).
- **H7** Heart rate is higher in PVR *vs* S (**H7P**) and higher in AVR *vs* PVR (**H7A**).
- **H8** Skin conductance is higher in PVR *vs* S (**H8P**) and higher in AVR *vs* PVR (**H8A**).

The last two hypotheses describe increasing effectiveness of the simulation in preparing users for harassment in terms of its perceived usefulness and elicited empathy:

- **H9** IMI Usefulness scores are higher in PVR *vs* S (**H9P**) and higher in AVR *vs* PVR (**H9A**).
- **H10** Empathy is higher in PVR *vs* S (**H10P**) and higher in AVR *vs* PVR (**H10A**).

5.1.3 Apparatus. In the screen condition, participants watched the video on a 24-inch screen with headphones while seated at a desk. In the VR conditions, participants wore an Oculus Quest 2 mobile head-mounted display (HMD). The HMD allowed for hand tracking without controllers, facilitating placement of skin conductance sensors on participants' fingers.

5.1.4 Procedure. After informed consent, participants underwent a screening process which had been developed with a harassment expert from the university ethics committee. The exclusion criteria were a history of extreme harassment, history of neurological disease, use of medication for psychological or emotional issues, epilepsy or use of medical devices (e.g. heart pump). The Trauma Screening Questionnaire [93] was also administered which asked whether participants had been affected at least twice in the previous week by an outcome of past harassment experience. This questionnaire contained 10 items such as whether participants had dreams or recollections about the event which could trigger negative emotions. More than four answers of 'yes' excluded the participant. Participants who passed the screening process completed a demographics questionnaire and baseline measures for Empathy, PANAS and STAI. Participants were given privacy to put on the heart rate sensor chest-strap, and then the skin conductance electrodes were attached to their fingers and the sensor unit attached to their wrist.

Next, participants started their first experimental condition. In the VR conditions (PVR and AVR), they were fitted with the HMD which at first displayed the Unity default infinite horizon to give participants time to adjust the HMD for comfort before the harassment scenario started. In the AVR condition, participants were encouraged to look at their hands and familiarise themselves with the tracking at the start of the experience. They were told that they could use their voice and hands to influence the situation if they wished. After each condition, participants removed the HMD, completed the empathy, PANAS, STAI, presence, agency and IMI questionnaires, and took a short break.

After the three conditions, participants completed the final, post-experiment questionnaire, were debriefed and rewarded for their participation with a £10 shopping voucher. One day after their participation, participants were contacted by the experimenter via email to check that participation in the study had not had adverse psychological consequences.

5.1.5 Participants. We recruited 44 participants (24 males, 20 females), aged 18-45 ($M = 26.86, SD = 6.26$), through social media and word of mouth. Participants generally had only a little experience with VR ($M = 3.65, SD = 2.84$ on a 10 point scale), considerable gaming experience ($M = 7.00, SD = 2.30$ on a 10 point scale), and high baseline levels of empathy ($M = 4.04, SD = 0.55$ on a 5 point scale). All participants had normal or corrected to normal vision and normal hearing.

5.2 Results

The results are illustrated in Figure 2. The error bars in the plots show 95% confidence intervals of the mean. The hypotheses were tested using paired, one-tailed t-tests. A power analysis using G*Power 3.1 [28] indicates that these t-tests were able to detect medium effects (Cohen's d = 0.381) at $\alpha = .05$ with a power of 0.8. In the absence of a hypothesis, we analysed differences across conditions S, PVR and AVR and effects of gender using two-way repeated-measures ANOVAs, with condition as within-participant factor and gender as between-participant factor. If a Mauchly test indicated a violation of sphericity, a Huynh-Feldt correction was applied [57]. If an ANOVA showed a significant main effect, pairwise t-tests with Bonferroni-Holm *post hoc* corrections were conducted, to control for multiple comparisons while maximising power [4, 13]. We provide effect sizes for all pairwise comparisons using Cohen's *d* to investigate the magnitude of the observed differences [46].

5.2.1 Manipulation Check. We performed pairwise t-tests to check whether the conditions S, PVR and AVR were able to elicit increased perceived agency and were able to elicit negative emotions, which are basic assumptions of this study. Agency was significantly greater in PVR compared to S ($t = -3.961, p < .001^{**}, d = -.597$), and greater in AVR compared to PVR ($t = -7.435, p < .001^{**}, d = -1.121$) (Figure 2d), indicating that the conditions succeeded in manipulating the levels of felt agency. PANAS scores were significantly greater than B ($M = 1.20, SD = 0.556$) in S ($t = -6.271, p < .001^{**}, d = -.992$), PVR ($t = -1.711, p < .001^{**}, d = -10.819$) and AVR ($t = -11.784, p < .001^{**}, d = -1.863$). IMI Pressure/Tension scores were higher than B ($M = 2.11, SD = 0.690$) in PVR ($t = -3.351, p = .007^{*}, d = -.537$) and in AVR ($t = -.469, p = .017^{*}, d = -.469$). IMI Pressure/Tension in S was not significantly different from B ($p = 1.000$). This indicates that the simulation conditions were successful in inducing emotional stress.

5.2.2 Presence. For Physical Presence (Figure 2a), the main effect of condition was significant ($F(1, 1.707) = 138.758, p < .001^{**}, eta_p^2 = .768$) but not of gender ($F(1, 42) = 1.011, p = .321$). PVR was significantly greater than S ($t = -12.961, p < .001^{**}, d = -1.954$) and AVR was significantly greater than PVR ($t = -2.600, p = .011^{*}, d = -0.393$), therefore, we accept **H1P** and **H1A**. For Social Presence (Figure 2b), the main effect of condition was significant ($F(1, 1.736) = 85.940, p < .001^{**}, eta_p^2 = .672$), but not of gender ($F(1, 42) = 0.034, p = .855$). PVR was significantly greater than S ($t = -10.986, p < .001^{**}, d = -1.656$) and also AVR was significantly greater than S ($t = -11.504, p < .001^{**}, d = -1.734$). However, AVR was not significantly greater than PVR ($p = .606$) (**H2A**), therefore, we accept **H2P** but not **H2A**. For Self Presence (Figure 2c), the main effect of condition was significant ($F(1, 2) = 98.727, p < .001^{**}, eta_p^2 = .702$), but not of gender ($F(1, 42) = 1.270, p = .266$). PVR was significantly greater than S ($t = -7.124, p < .001^{**}, d = -1.074$) and AVR was significantly greater than PVR ($t = -7.031, p < .001^{**}, d = -1.060$), therefore, we accept **H3P** and **H3A**.

5.2.3 Emotion. For PANAS (Figure 2e), the main effect of condition was significant ($F(1, 2.052) = 7.948, p < .001^{**}, eta_p^2 = .401$), but not of gender ($F(1, 42) = 1.269, p = .266$). PVR was significantly greater than S ($t = -5.622, p < .001^{**}, d = -.848$) and AVR was also significantly greater than S ($t = -6.754, p < .001^{**}, d = -1.018$). However, AVR was not significantly greater than PVR ($p = .261$), therefore, we accept **H4P** but not **H4A**. For STAI (Figure 2f), the main effect of condition was significant ($F(1, 2.076) = 20.766, p < .001^{**}, eta_p^2 = .309$), but not of gender ($F(1, 42) = 2.022, p = .162$). PVR was significantly greater than S ($t = -5.546, p < .001^{**}, d = -.836$) and AVR was significantly greater than S ($t = -5.179, p < .001^{**}, d = -.781$). However, AVR was not significantly greater than PVR ($p = .715$), therefore, we accept **H5P** but not **H5A**. For Pressure/Tension (Figure 2h), the main effect of condition was significant ($F(1, 2) = 23.796, p < .001^{**}, eta_p^2 = .352$) and so was gender ($F(1, 42) = 4.319, p = .044^{*}, eta_p^2 = .093$). PVR was significantly greater than S ($t = -4.716, p < .001^{**}, d = -.711$) and also AVR was significantly greater than S ($t = -6.696, p < .001^{**}, d = -1.009$). However, AVR was not significantly greater than PVR ($p = .51$), therefore, we accept **H6P** but not **H6A**. For HR (Figure 2j), the main effect of condition was significant ($F(1, 1.484) = 48.714, p < .001^{**}, eta_p^2 = .537$), but not of gender ($F(1, 42) = 0.941, p = .337$). An interaction effect was found between condition and gender ($F(1, 14.84) = 4.027, p = .034^{*}, eta_p^2 = .087$). PVR was significantly greater than S ($t = -5.811, p < .001^{**}, d = -1.027$) and also AVR was significantly greater than S ($t = -7.669, p < .001^{**}, d = -1.356$). However, AVR was not significantly greater than PVR ($p = .068$), therefore, we accept **H7P** but not **H7A**. For skin conductance (Figure 2k), the main effect of condition was significant ($F(1, 2) = 8.261, p = .002^{*}, eta_p^2 = .141$), but not of gender ($F(1, 42) = 3.970, p = .053$). AVR was significantly greater than S ($t = -2.648, p = .031^{*}, d = -.468$). PVR was not significantly greater than S ($p = .283$) and AVR was not significantly greater than PVR ($p = .245$), therefore, we reject both **H8P** and **H8A**.

5.2.4 Usefulness. For the IMI Value/Usefulness measure (Figure 2i), the main effect of condition was significant ($F(1, 2) = 21.845, p < .001^{**}, eta_p^2 = .090$), but not of gender ($F(1, 42) = 0.579, p = .451$). PVR was significantly greater than S ($t = -5.253, p < .001^{**}, d = -.792$) and also AVR was significantly greater than S ($t = -6.248, p < .001^{**}, d = -.942$). However, AVR was not significantly greater than PVR ($p = .322$), therefore, we accept **H9P** but not **H9A**. For Empathy (Figure 2g), the main effect of condition was significant ($F(1, 1.977) = 3.564, p < .001^{**}, eta_p^2 = .078$) and so was the effect for gender ($F(1, 42) = 18.722, p < .001^{**}, eta_p^2 = .308$). PVR was

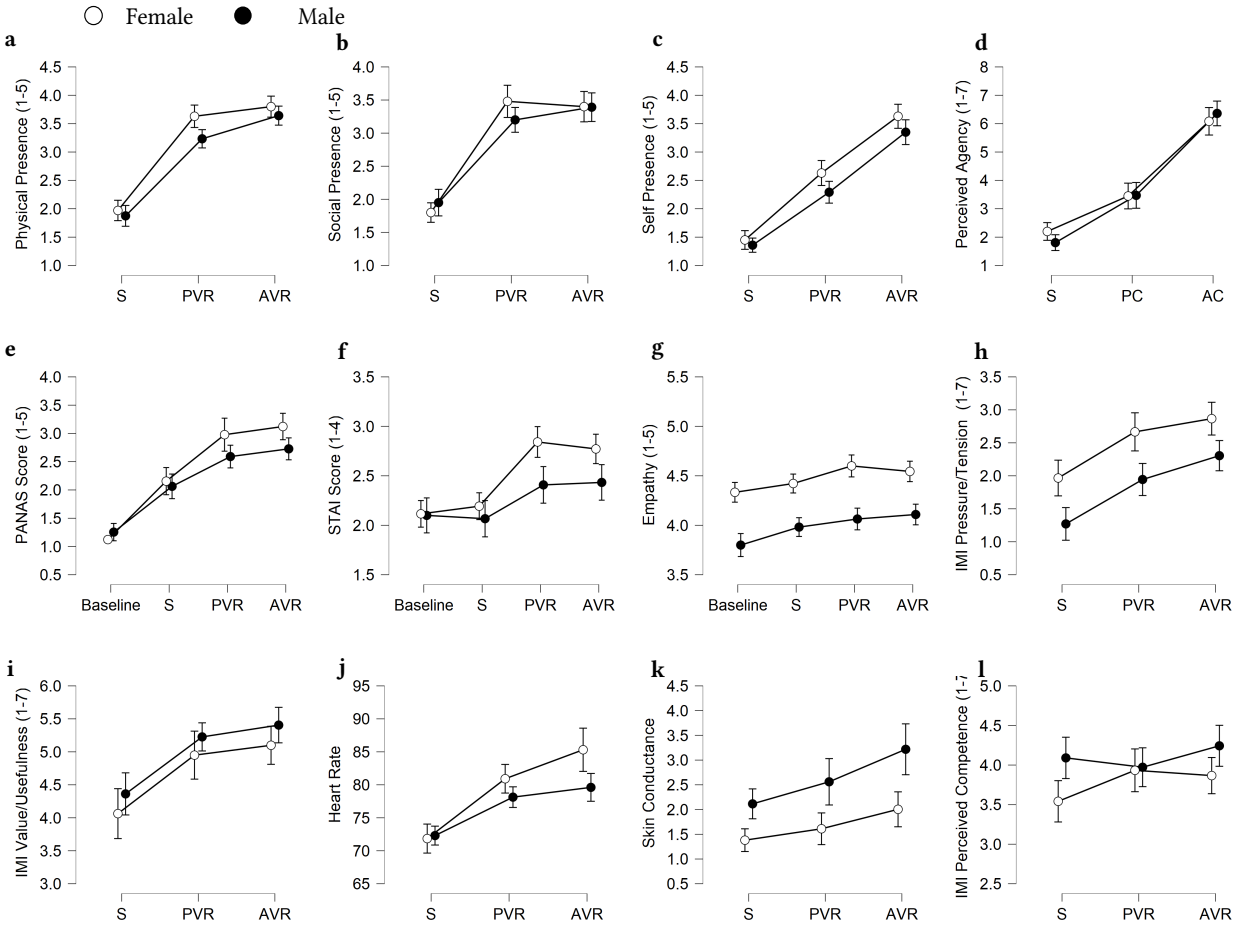

Figure 2: Presence, emotion and effectiveness for baseline (if available) and in the screen (S), passive VR (PVR), and active VR (AVR) conditions.

Table 1: Summary of results for Study 3 (mean ± std. dev.). Asterisks indicate significant differences between Passive VR and Screen, and Active VR and Passive VR, respectively.

Variable (Hypothesis)	Screen	Passive VR	Active VR
Agency	1.99±1.34	**3.46±2.11	**4.30±0.53
Physical Presence (H1)	1.92±0.08	**3.41±0.84	*3.71±0.08
Social Presence (H2)	1.88±0.08	**3.33±0.99	3.39±1.03
Self Presence (H3)	1.40±0.06	**2.45±0.96	**3.48±1.01
PANAS (H4)	2.07±0.98	**2.73±1.12	2.86±0.96
STAI (H5)	2.08±.071	**2.54±.082	2.51±0.78
IMI Press./Tension (H6)	1.59±1.25	**2.27±1.27	2.56±1.14
Heart Rate (H7)	71.96±1.18	**79.03±8.24	81.30±10.32
Skin Conductance (H8)	1.71±1.40	*1.99±1.92	2.39±2.06
IMI Value/Useful. (H9)	4.23±1.61	**5.10±1.33	5.26±1.30
Empathy (H10)	4.17±0.50	*4.32±0.55	4.30±0.53

significantly greater than S ($t = -2.550, p = .038^{*}, d = -.384$). AVR, however, was not significantly greater than S ($p = .089$). PVR and AVR were not significantly different either ($p = .611$), therefore, we accept **H10P**) but not **H10A**. Scores for empathy differed between male and female participants in S ($t = 3.848, p = .003^{*}$), PVR ($t = 4.027, p = .002^{*}$) and AVR ($t = 3.729, p = .004^{*}$). For Perceived Competence (Figure 2l), the main effect of condition was not significant ($F(1, 2) = 1.018, p = .368$), and neither was gender ($F(1, 42) = 1.094, p = .302$).

5.2.5 Realism Rankings & Qualitative Results. When ranking the three conditions in terms of realism, 38 of our 44 participants ranked AVR as the most realistic, with six saying PVR was most realistic and none for S. All but two participants thought the S condition was least realistic. Overall, responses mentioned that the visual and auditory representation of the simulation had a significant influence on perceiving the virtual scenario as real. Experiencing the harassment simulation on a screen did not have a strong immersive effect, with participants mentioning their awareness of the testing

room. Although the term was not explicitly used, most participants mentioned feelings of physical presence in the VR conditions. The ability to view the surroundings by moving their heads in any direction in VR induced a sense of being enveloped in the virtual space and the scenario (*"The VR simulations felt more immersive due to how much physically closer the objects and people felt"*).

Due to the resemblance of the virtual harasser's behaviour and attitude to a real life scenario, the persona of the harasser evoked discomfort in participants, indicating that participants felt present in the virtual environment (*"In VR, it was difficult not to feel like I was there, and it was very uncomfortable at times"*). The ability to see their own hands in the interactive VR allowed participants to embody the victim, which was not possible in the screen condition (*"I felt more like I embodied the victim because I could see my hands – I didn't have any sense that my real hands were separate from the virtual hands, it's like they had fused"*). Interaction with the harasser via speaking or with their hands gave participants a sense of agency (*"Being able to interact with my hands and pushing the harasser away helped greatly in terms of making me feel involved and also having sense of control"*). Nevertheless, in some instances, participants felt the interaction to be forced and unrealistic (*"I think the interactive VR harassment experience didn't seem that realistic because it is forced interaction, rather than my normal behaviour. Ordinarily, how I would respond in that situation would be to walk away which I couldn't do"*). Participants discussed that the feeling of a forced interaction was also due to the lack of threat in the environment (*"it was difficult to get immersed in a 'threatening' situation that lacked any real threat. I feel this would be one of the main factors that would influence behaviour in a 'real' scenario"*).

5.2.6 Usefulness Qualitative Results. Most participants were positive about the effectiveness of the tool in empowering people. They highlighted that the virtual environment would create a sense of safety and autonomy, enabling them to easily take steps necessary to cope with the psychological consequences of street harassment (*"I think it could teach people to feel more powerful in situations of street harassment. When I have been harassed previously, I was so shocked by what was happening that I felt like I couldn't react. Familiarising oneself with this kind of scenario could help people to feel more in control of their response to the harassment"*). Participants also noted that experiencing in VR what it can be like to be harassed and put in an uncomfortable and threatening situation enhanced empathy toward victims of street harassment. Not all participants had experienced street harassment in real life. Hence, introducing the perspective of the victim did produce positive effects, such as greater appreciation of the psychological consequences of street harassment, increased likelihood to intervene as a bystander, or changing their attitudes and behaviours to prevent becoming a harasser themselves (*"Perhaps making people ambivalent to street harassment have some sense of what it feels like to be in that situation, how persistent people can be and how intimidating it feels when they get up in your space and won't go away. This might make them more likely to intervene in situations, regulate their friends' behaviours or even not do it themselves"*).

5.2.7 Ethics Qualitative Results. Participants disagreed that the experience was unethical, so long as there is a screening procedure and users are given prior information regarding the VR simulation and potential risks. One important aspect raised by many of the participants was that the VR simulation creates a safe environment and gives them agency to end the simulation at any point. The sense of safety may make it easier for victims to cope with their past traumas at their own pace. Additionally, given that street harassment can occur anywhere and at any time, individuals may panic and not know how to deal with the situation. The safe environment created by the simulation can give individuals time and peace of mind to learn how to manage and prepare for such anxiety-provoking circumstances (*"I think this simulation is ethical because it gives individuals a chance to work through any emotions experienced and how to deal with such situations. In these situations it's hard to know how you will react and as it comes up randomly it's hard to prepare. Knowing you can step out of the situation immediately makes it a slightly easier way to practise dealing with the situation. It feels more realistic than an acted out scenario would feel in a standard room without the individual being actually in danger of harassment"*).

None of the participants saw the VR tool as a form of victim blaming. Rather, the simulation was perceived as depicting a real life event that is prevalent in society (*"I do not think that this VR tool appears as a form of victim blaming, rather it points towards the current trends that exist within harassment. Some of the comments that were said by the perpetrator included how well the victim was dressed and looked, as well as the fact that it was okay to say such things at 9pm if someone was dressed like that. These comments are quite realistic in my opinion and do an effective job of portraying such real life events. I do not think that the VR tool normalises such comments, but highlights them instead. Hence, it plays an educative function as it creates awareness"*, *"I think if it is intended for use as a preparation for people facing street harassment I think it would be useful, as women or men need all the preparation in case of encountering a scenario such as this"*).

6 DISCUSSION

6.1 How can we design a realistic simulation of street harassment? (RQ1)

Very little research has investigated which design elements are most important in the creation of harassment simulations [39, 40, 85], with no research focusing on street harassment in particular. Existing training tools did not base their designs on published empirical scientific results [1–3]. Our results indicate that several design features are fundamental to an effective simulation of street harassment. First, ecological validity was regarded by both users and experts as most important for the tool to be effective in inducing presence and realistic emotions. It is therefore crucial to base such tools on real harassment scenarios. Similarly, the quality of audiovisual content representing the harasser was found to be important to 'bring him to life', including the script, voice acting, avatar model and animations. Although some emotional responses can be achieved when using only a 2D screen, Study 3 shows that it is much more effective to use immersive VR. This suggests that future tools should design scenarios for common consumer HMDs to optimise availability and effectiveness. Our hypotheses, which were almost all accepted, clearly indicate a progression from screen to passive VR to active VR with increasing presence and emotional response. Although the effect levels off somewhat from passive

to active VR, it is clear that interactive features increase realism further by actively engaging the user in the experience. It is worth noting that the interactions we evaluated were quite simple and yet showed clear effects, suggesting that it is not necessary to spend a lot of resources and effort on complex interaction design.

The implementation of the two interactive features increased physical presence and self-presence as expected but social presence was not higher in AVR compared to PVR, which is in contrast to previous findings [69]. This could be due to the interaction methods or the reactions of the avatar not being realistic enough. Again as expected, perception of agency increased with the presence of the interactive features. However, none of the PANAS, STAI, IMI Pressure/Tension or Empathy measures was significantly different in AVR compared to PVR, which suggests that emotions were not much affected by the increase in agency, as was also reflected in the physiological measures. This dissociation between agency and emotion has been observed before in the case of fear [38]. Lastly, the IMI Usefulness scores did not see an improvement with the addition of agency, which suggests that, by this measure at least, users did not perceive the value of the interactive features and the increased agency they offered.

6.2 What are the ethical issues surrounding simulations of street harassment? (RQ2)

Domain experts raised several ethical issues around the simulation of street harassment. First, they noted that for some users, experiencing a harassment scenario could trigger severe negative emotions from past extreme harassment experiences. However, they believed that this risk can be effectively mitigated through a screening process and the presence of professionals during the simulation. In this respect, they suggested design features that would enable users to mitigate risks: the simulation should inform users about what to expect in the experience and clearly warn them about possible negative effects; it should tell them how to get support if needed, e.g. through phone help lines; and it should enable users to quickly and easily stop the experience at any time. The most appropriate setting for such a simulation is in expert-led anti-harassment activities such as Empowerment Self Defence courses.

Harassment simulations may be perceived as victim blaming, especially if they are framed as 'training' for potential victims. The experts believed that this could be addressed by clarifying their context and purpose. It should be made explicit that the responsibility and blame for harassment lies entirely with the harasser, ideally as part of the simulation's design. Finally, simulations can never represent the full spectrum of possible harassment scenarios in the real world. Therefore, they should not give users the impression that they are fully representative of harassment, that all behaviours other than those shown would be acceptable, or that particular behavioural responses would be appropriate in all situations. The development and real world application of behavioural strategies should be guided by domain experts.

6.3 How effective is a simulation of street harassment in eliciting realistic emotions? (RQ3)

Evaluations of tools intended to prepare people for harassment generally rely on plausible indicators of effectiveness, as it would be unethical intentionally to expose users to harassment in the real world. Hence, we collected evidence from experts (Study 2) and users (Study 3) that estimate the simulation's likely effectiveness based on qualitative and quantitative measures. Our findings suggest that the simulation can be effective: most domain experts believed that it would be useful and several asked us to make the simulation available for their professional use. This was supported by users' comments and the results from well validated scales and physiological measurements.

According to SIT theory, the simulation's effectiveness in eliciting realistic emotions is likely to translate into increased resilience and ability to cope [65]. Domain experts regarded the simulation as a 'strong tool' that would make preparedness training both more practical and more widely available. Furthermore, improvements in measures of intrinsic motivation such as IMI Perceived Competence and Usefulness/Value are often associated with increased performance [12] and positive psychological and behavioural developments [75]. Although this has not been validated for SIT, users' comments indicate that the simulation generally increased their sense of empowerment and preparedness. Despite the elicited negative emotions, they felt safe and in control, similar to previous work on the use of VR in relation to harassment [39, 40, 52, 85]. Finally, related work also suggests that the simulation's ability to increase empathy for harassment victims could translate into positive behaviour change [64, 91]. Consonant with this finding, participants reported that the simulation had increased their understanding of how victims of street harassment feel, and suggested that it could make people more likely to intervene as bystanders and change attitudes and behaviours, thereby preventing them from exhibiting harassment behaviours themselves.

An interesting observation is the significantly higher empathy reported by women in all conditions. Women also scored higher on the IMI Pressure/Tension scale, suggesting that they felt more pressure and discomfort. These results are not surprising given that the overwhelming majority of harassment victims are women [86]. These results are in line with findings on gender differences in perceptions of harassment, which indicate that harassment is generally perceived as more serious and threatening by women than by men [25, 53, 74]. The lack of notable gender effects for any of the other measures indicates that the tool can be effective in eliciting presence and negative emotions across genders.

6.4 Limitations

The domain experts emphasised that, while our selected scenario was appropriate, harassment can take many forms which cannot all be represented in a single simulation. They would like a greater range and variety of virtual harassment scenarios to be implemented. One ESD instructor stated that successful training is usually achieved with 18-20 scenarios. Experts also suggested that the tool should be adaptable to individual characteristics such as different demographics, and that control over the intensity of harassment

could be useful. Furthermore, some users mentioned that the range of interactions implemented was too limited to allow them to react in the way they naturally would, such as walking away from the harasser. This could, in part, explain the lack of differences between the PVR and AVR conditions for most of the outcome variables. Still, it should be noted that the interactive features implemented did not jeopardise the perceived realism of the situation, interaction or harasser (as shown by the results presented in Appendix 3).

No VR simulation could replicate every possible street harassment scenario and offer all possible interaction methods. However, having demonstrated the potential of this approach with a common scenario, we can work with domain experts and others to extend the range of simulated scenarios. Similarly, additional interaction methods can readily be implemented as the simulation tool is further refined, in order to further enhance presence and potentially other variables.

Another limitation was the use in Study 1 of videos to present features of the VR experience, instead of an HMD. Similarly, the short duration of the main video (2 minutes) may not have allowed levels of presence to be achieved similar to the 5 minute long main experience presented in Study 3. Nonetheless, Study 1 served to inform further design decisions and more general guidelines; and it was not intended to be a thorough assessment of the experience which was then in an incipient phase. This assessment was reserved for Study 3.

Finally, our participants were mostly young adults (age $M = 26.86, SD = 6.26$), while harassment is experienced by people of all ages [68]. We therefore do not know how far our results can be generalised to different age groups, or the extent to which our simulation can be effective with older adults, who have been shown to rate VR interventions lower on usability compared to younger people [45]. These limitations are all promising directions for future work.

6.5 Future Work

Throughout Studies 2 and 3 it became apparent that our simulation could also serve purposes other than the initial intended one of psychologically preparing people for experiencing street harassment. ESD instructors felt the tool would be effective if used alongside ESD training, as it could be used to demonstrate common scenarios so that appropriate training could be provided by a professional. In the time of a pandemic, ESD and anti-harassment training has been provided remotely; when paired with such remote training, VR simulations could make training with realistic scenarios more accessible to people who are not able to attend or not comfortable attending lessons in person. Moreover, both experts and users suggested that the simulation could be effectively applied to other ways of combating street harassment, such as using it as an awareness raising tool for potential harassers and for bystander training [31, 64]. Yet another use that was proposed was in practising de-escalation techniques for specific types of scenario, if users were able to influence the situation with more interactive features tailored around de-escalation. In summary, VR simulations of harassment can serve multiple purposes and can likely be integrated into existing practices that aim to combat street harassment, enhancing their accessibility and effectiveness.

While we have evaluated the effects of our design choices on realism only in the context of a street harassment scenario, it is plausible that they generalise to some degree to other scenarios. There are many applications that aim to induce negative emotions, e.g. for other types of stress inoculation training [58, 65, 96] and exposure therapy [11, 66, 95]. Our results may inform the design of other scenarios, e.g. by illustrating the benefits of immersive VR and interactivity, in particular for those that have a strong interpersonal component such as in the treatment of social anxiety [14, 27].

It is important to acknowledge that the impact of street harassment is much wider than we could cover in the present study. For example, many gender diverse people suffer street harassment disproportionately more than others [21, 62]. As noted in our expert interviews in Study 2, the scenario we have used in this study is only one of many possible scenarios. While many of our findings on realism in simulation design (RQ1), ethical issues in simulating street harassment (RQ2) and the effectiveness of VR in eliciting realistic emotions (RQ3) are likely to apply in other scenarios and with different people, more work on *in virtuo* simulations of harassment is needed, including diversity of both victims and harassers and consideration of the effectiveness of different behaviours in different scenarios.

7 CONCLUSION

We co-designed a VR simulation of a street harassment situation with experts and end-users, investigating VR design choices and their effects on realism and effectiveness, as well as the surrounding ethical issues. Our results are intended to help VR designers and other stakeholders to realise the potential of immersive technology in combating street harassment, which has been largely unexplored and still has many promising directions of future work. In summary:

(1) Ecologically valid scenarios with high quality audiovisual representations of the harasser can effectively immerse users and elicit a realistic emotional response.
(2) Simulating harassment in VR has inherent ethical risks such as emotional triggering and potential perception of victim blaming, which can and should be addressed in the design and use of the simulation.
(3) VR harassment simulations can serve multiple purposes including both increasing psychological preparedness for being harassed and enhancing awareness and empathy in potential harassers, and can extend and complement existing methods of combating harassment.

ACKNOWLEDGMENTS

This work was supported and partly funded by the Centre for the Analysis of Motion, Entertainment Research and Applications (CAMERA 2.0; EP/T022523/1) at the University of Bath.

REFERENCES

[1] [n.d.]. *EmpowHER: Self-Defense Training.* https://www.oculus.com/experiences/rift/2923737841073277/?locale=en_GB
[2] [n.d.]. *Regatta VR.* https://regattavr.com/
[3] [n.d.]. *Vantage Point.* https://www.tryvantagepoint.com/
[4] Mikel Aickin and Helen Gensler. 1996. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. *American journal of public health* 86, 5 (1996), 726–728.

[5] APPG for UN Women. 2021. Prevalence and reporting of sexual harassment in UK public spaces. (2021).

[6] Ali Azarbarzin, Michele Ostrowski, Patrick Hanly, and Magdy Younes. 2014. Relationship between arousal intensity and heart rate response to arousal. *Sleep* 37, 4 (2014), 645–653.

[7] Rosa María Baños, Cristina Botella, Mariano Alcañiz, Víctor Liaño, Belén Guerrero, and Beatriz Rey. 2004. Immersion and emotion: their impact on the sense of presence. *Cyberpsychology & behavior* 7, 6 (2004), 734–741.

[8] Alan S Bellack. 1983. Recurrent problems in the behavioral assessment of social skill. *Behaviour Research and Therapy* 21, 1 (1983), 29–41.

[9] Michelle CD Blaxall, Barry S Parsonsona, and Neville R Robertson. 1993. The development and evaluation of a sexual harassment contact person training package. *Behavior modification* 17, 2 (1993), 148–163.

[10] Cynthia Grant Bowman. 1993. Street Harassment and the Informal Ghettoization of Women. *Harvard Law Review* 106, 3 (1993), 517–580. http://www.jstor.org/stable/1341656

[11] Emily Carl, Aliza T Stein, Andrew Levihn-Coon, Jamie R Pogue, Barbara Rothbaum, Paul Emmelkamp, Gordon JG Asmundson, Per Carlbring, and Mark B Powers. 2019. Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of anxiety disorders* 61 (2019), 27–36.

[12] Christopher P Cerasoli, Jessica M Nicklin, and Michael T Ford. 2014. Intrinsic motivation and extrinsic incentives jointly predict performance: a 40-year meta-analysis. *Psychological bulletin* 140, 4 (2014), 980.

[13] Shi-Yi Chen, Zhe Feng, and Xiaolian Yi. 2017. A general introduction to adjustment for multiple comparisons. *Journal of thoracic disease* 9, 6 (2017), 1725.

[14] Rachel K Chesham, John M Malouff, and Nicola S Schutte. 2018. Meta-analysis of the efficacy of virtual reality exposure therapy for social anxiety. *Behaviour Change* 35, 3 (2018), 152–166.

[15] Joon Hao Chuah, Andrew Robb, Casey White, Adam Wendling, Samsun Lampotang, Regis Kopper, and Benjamin Lok. 2012. Increasing agent physicality to raise social presence and elicit realistic behavior. In *2012 IEEE Virtual Reality Workshops (VRW)*. IEEE, 19–22.

[16] M Meghan Davidson, Michael S Butchko, Krista Robbins, Lindsey W Sherd, and Sarah J Gervais. 2016. The mediating role of perceived safety on street harassment and anxiety. *Psychology of Violence* 6, 4 (2016), 553.

[17] Edward L Deci. 86. Ryan. RM (1985). Intrinsic motivation and self-determination in human behavior. *New York and London: Plenum* 86 (86), 18–19.

[18] Maria DelGreco and John Christensen. 2020. Effects of street harassment on anxiety, depression, and sleep quality of college women. *Sex Roles* 82, 7 (2020), 473–481.

[19] Micaela Di Leonardo. 1981. The political economy of street harassment. *Aegis* (1981), 51–57.

[20] Julia Diemer, Georg W Alpers, Henrik M Peperkorn, Youssef Shiban, and Andreas Mühlberger. 2015. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology* 6 (2015), 26.

[21] Petra L Doan. 2010. The tyranny of gendered spaces–reflections from beyond the gender dichotomy. *Gender, Place & Culture* 17, 5 (2010), 635–654.

[22] Anders Drachen, Lennart E Nacke, Georgios Yannakakis, and Anja Lee Pedersen. 2010. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*. 49–54.

[23] Tracey L Earle, Wolfgang Linden, and Joanne Weinberg. 1999. Differential effects of harassment on cardiovascular and salivary cortisol stress reactivity and recovery in women and men. *Journal of psychosomatic research* 46, 2 (1999), 125–141.

[24] Darragh Egan, Sean Brennan, John Barrett, Yuansong Qiao, Christian Timmerer, and Niall Murray. 2016. An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments. In *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 1–6.

[25] John O Ekore. 2012. Gender differences in perception of sexual harassment among university students. *Gender and Behaviour* 10, 1 (2012), 4358–4369.

[26] Mai ElSherief and Elizabeth Belding. 2015. The Urban Characteristics of Street Harassment: A First Look. In *Proceedings of the 1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics*. 28–31.

[27] Paul MG Emmelkamp, Katharina Meyerbröker, and Nexhmedin Morina. 2020. Virtual reality therapy in social anxiety disorder. *Current psychiatry reports* 22, 7 (2020).

[28] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191.

[29] Barbara L Fredrickson and Tomi-Ann Roberts. 1997. Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of women quarterly* 21, 2 (1997), 173–206.

[30] Peter Fromberger, Sabrina Meyer, Christina Kempf, Kirsten Jordan, and Jürgen L Müller. 2015. Virtual viewing time: the relationship between presence and sexual interest in androphilic and gynephilic men. *PloS one* 10, 5 (2015), e0127156.

[31] Sarah Garcia, Soumya Joseph Abraham, and Marvin Andujar. 2021. Exploring Perceptions of Bystander Intervention Training using Virtual Reality. In *ACM International Conference on Interactive Media Experiences*. 253–257.

[32] Carol Brooks Gardner. 1995. *Passing by: Gender and public harassment*. Univ of California Press.

[33] Tom Geller. 2008. Overcoming the uncanny valley. *IEEE computer graphics and applications* 28, 4 (2008), 11–17.

[34] Sharon Goulds, Sophie Tanner, Nicole Kalms, and Gill Matthewson. 2018. *Unsafe in the City - The Everyday Experiences of Girls and Young Women: The State of the World's Girls 2018*. Plan International. MDQLTY - F.

[35] Wendy Hadley, Christopher Houck, Larry K Brown, Josh S Spitalnick, Mirtha Ferrer, and David Barker. 2019. Moving beyond role-play: evaluating the use of virtual reality to teach emotion regulation for the prevention of adolescent risk behavior within a randomized pilot trial. *Journal of pediatric psychology* 44, 4 (2019), 425–435.

[36] André Helgert, Sabrina C Eimler, and Alexander Arntz. 2021. Stop Catcalling-A Virtual Environment Educating Against Street Harassment. In *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 419–421.

[37] Kathryn J Holland, Verónica Caridad Rabelo, Amber M Gustafson, Rita C Seabrook, and Lilia M Cortina. 2016. Sexual harassment against men: Examining the roles of feminist activism, sexuality, and organizational context. *Psychology of Men & Masculinity* 17, 1 (2016), 17.

[38] Crescent Jicol, Chun Hin Wan, Benjamin Doling, Caitlin H Illingworth, Jinha Yoon, Charlotte Headey, Christof Lutteroth, Michael J Proulx, Karin Petrini, and Eamonn O'Neill. 2021. Effects of Emotion and Agency on Presence in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[39] Ernest N Jouriles, Renee McDonald, Antje Kullowatz, David Rosenfield, Gabriella S Gomez, and Anthony Cuevas. 2009. Can virtual reality increase the realism of role plays used to teach college women sexual coercion and rape-resistance skills? *Behavior therapy* 40, 4 (2009), 337–345.

[40] Ernest N Jouriles, Lorelei Simpson Rowe, Renee McDonald, Cora G Platt, and Gabriella S Gomez. 2011. Assessing women's responses to sexual threat: Validity of a virtual role-play procedure. *Behavior therapy* 42, 3 (2011), 475–484.

[41] Mike Just, Hasmeet Chandok, Raghav Sampangi, Kirstie Hawkey, Alette Willis, JeyaBalaji Samuthiravelu, Dilpreet Gill, and Michael Altair. 2019. Personal Safety App Effectiveness. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.

[42] Naveena Karusala and Neha Kumar. 2017. Women's safety in public spaces: Examining the efficacy of panic buttons in New Delhi. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3340–3351.

[43] Holly Kearl. 2014. Unsafe and harassed in public spaces: A national street harassment report. (2014).

[44] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (2012), 373–387.

[45] Panagiotis Kourtesis, Simona Collina, Leonidas AA Doumas, and Sarah E MacPherson. 2019. Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology. *Frontiers in human neuroscience* 13 (2019), 417.

[46] Daniël Lakens. 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology* 4 (2013), 863.

[47] Po-Hung Lin and Jun-Hao Chen. 2020. Effects of video type, display technique, and ambient illumination on visual and physiological performance. *Journal of Ambient Intelligence and Humanized Computing* (2020), 1–8.

[48] Beth Livingston, M Grillo, and R Paluch. 2014. Cornell international survey on street harassment.

[49] Roberto Lloréns, Enrique Noé, Carolina Colomer, and Mariano Alcañiz. 2015. Effectiveness, usability, and cost-benefit of a virtual reality–based telerehabilitation program for balance recovery after stroke: A randomized controlled trial. *Archives of physical medicine and rehabilitation* 96, 3 (2015), 418–425.

[50] Matthew Lombard, Theresa B Ditton, and Lisa Weinstein. 2009. Measuring presence: the temple presence inventory. In *Proceedings of the 12th annual international workshop on presence*. 1–15.

[51] Claudie Loranger and Stéphane Bouchard. 2017. Validating a virtual environment for sexual assault victims. *Journal of traumatic stress* 30, 2 (2017), 157–165.

[52] Laura Loucks, Carly Yasinski, Seth D Norrholm, Jessica Maples-Keller, Loren Post, Liza Zwiebach, Devika Fiorillo, Megan Goodlin, Tanja Jovanovic, Albert A Rizzo, et al. 2019. You can do that?!: Feasibility of virtual reality exposure therapy in the treatment of PTSD due to military sexual trauma. *Journal of anxiety disorders* 61 (2019), 55–63.

[53] Manish Madan and Mahesh K Nalla. 2016. Sexual harassment in public spaces: Examining gender differences in perceived seriousness and victimization. *International criminal justice review* 26, 2 (2016), 80–97.

[54] Moira Maguire and Brid Delahunt. 2017. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education* 9, 3 (2017).

[55] Guido Makransky, Lau Lilleholt, and Anders Aaby. 2017. Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior* 72 (2017), 276–285.

[56] Abigail A Marsh and Robert James R Blair. 2008. Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews* 32, 3 (2008), 454–465.

[57] Scott E Maxwell, Harold D Delaney, and Ken Kelley. 2017. *Designing experiments and analyzing data: A model comparison perspective.* Routledge.

[58] Donald Meichenbaum. 2017. Stress inoculation training: A preventative and treatment approach. In *The evolution of cognitive behavior therapy.* Routledge, 101–124.

[59] Miguel Melo, Sofia Sampaio, Luís Barbosa, José Vasconcelos-Raposo, and Maximino Bessa. 2016. The impact of different exposure times to 360 video experience on the sense of presence. In *2016 23rd Portuguese Meeting on Computer Graphics and Interaction (EPCGI).* IEEE, 1–5.

[60] Stanley Milgram. 1963. Behavioral study of obedience. *The Journal of abnormal and social psychology* 67, 4 (1963), 371.

[61] Mats Najström and Billy Jansson. 2007. Skin conductance responses as predictor of emotional responses to stressful life events. *Behaviour research and therapy* 45, 10 (2007), 2456–2463.

[62] Viviane K Namaste. 2013. *Genderbashing: Sexuality, Gender, and the Regulation of Public Space.* Routledge.

[63] Amie R Newins, Laura C Wilson, and Rebekah Z Kanefsky. 2020. What's in a label? The impact of media and sexual assault characteristics on survivor rape acknowledgment. *Journal of Traumatic Stress* (2020).

[64] Solène Neyret, Xavi Navarro, Alejandro Beacco, Ramon Oliva, Pierre Bourdin, Jose Valenzuela, Itxaso Barberia, and Mel Slater. 2020. An embodied perspective as a victim of sexual harassment in virtual reality reduces action conformity in a later Milgram obedience scenario. *Scientific reports* 10, 1 (2020), 1–18.

[65] Federica Pallavicini, Luca Argenton, Nicola Toniazzi, Luciana Aceti, and Fabrizia Mantovani. 2016. Virtual reality applications for stress management training in the military. *Aerospace medicine and human performance* 87, 12 (2016), 1021–1030.

[66] Mark B Powers and Paul MG Emmelkamp. 2008. Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of anxiety disorders* 22, 3 (2008), 561–569.

[67] Daniel J Rea, Denise Geiskkovitch, and James E Young. 2017. Wizard of awwws: Exploring psychological impact on the researchers in social HRI experiments. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* 21–29.

[68] Laura A Reese and Karen E Lindenberg. 2005. Gender, age, and sexual harassment. *Review of Public Personnel Administration* 25, 4 (2005), 325–352.

[69] Giuseppe Riva. 2002. The sociocognitive psychology of computer-mediated communication: The present and future of technology-based interactions. *Cyberpsychology & behavior* 5, 6 (2002), 581–598.

[70] Giuseppe Riva, Fabrizia Mantovani, Claret Samantha Capideville, Alessandra Preziosa, Francesca Morganti, Daniela Villani, Andrea Gaggioli, Cristina Botella, and Mariano Alcañiz. 2007. Affective interactions using virtual reality: the link between presence and emotions. *Cyberpsychology & behavior* 10, 1 (2007), 45–56.

[71] Albert Rizzo. 2005. Bravemind: virtual reality exposure therapy. *University of Southern California Institute for Creatrive Technologies* (2005).

[72] Albert 'Skip' Rizzo and Russell Shilling. 2017. Clinical virtual reality tools to advance the prevention, assessment, and treatment of PTSD. *European Journal of Psychotraumatology* 8, sup5 (2017), 1414560.

[73] Shane L Rogers, Ross Hollett, Yanqi R Li, and Craig P Speelman. 2020. An Evaluation of Virtual Reality Role-Play Experiences for Helping-Profession Courses. *Teaching of Psychology* (2020), 0098628320983231.

[74] Maria Rotundo, Dung-Hanh Nguyen, and Paul R Sackett. 2001. A meta-analytic review of gender differences in perceptions of sexual harassment. *Journal of applied psychology* 86, 5 (2001), 914.

[75] Richard M Ryan, Heather Patrick, Edward L Deci, and Geoffrey C Williams. 2008. Facilitating health behaviour change and its maintenance: Interventions based on self-determination theory. *The European health psychologist* 10, 1 (2008), 2–5.

[76] D Sapkaroski, M Mundy, and MR Dimmock. 2020. Virtual reality versus conventional clinical role-play for radiographic positioning training: a students' perception study. *Radiography* 26, 1 (2020), 57–62.

[77] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 266–281.

[78] Martijn J Schuemie, Peter Van Der Straaten, Merel Krijn, and Charles APG Van Der Mast. 2001. Research on presence in virtual reality: A survey. *CyberPsychology & Behavior* 4, 2 (2001), 183–201.

[79] Sofia Seinfeld, Minye Zhan, Marta Poyo-Solanas, Giulia Barsuola, Maarten Vaessen, Mel Slater, Maria V Sanchez-Vives, and Beatrice de Gelder. 2021. Being the victim of virtual abuse changes default mode network responses to emotional expressions. *cortex* 135 (2021), 268–284.

[80] Anil K Seth, Keisuke Suzuki, and Hugo D Critchley. 2012. An interoceptive predictive coding model of conscious presence. *Frontiers in psychology* 2 (2012), 395.

[81] Grant S Shields, Matthew A Sazma, and Andrew P Yonelinas. 2016. The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience & Biobehavioral Reviews* 68 (2016), 651–668.

[82] Yu Shu, Yen-Zhang Huang, Shu-Hsuan Chang, and Mu-Yen Chen. 2019. Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. *Virtual Reality* 23, 4 (2019), 437–446.

[83] Matthew J Smith, Emily J Ginger, Katherine Wright, Michael A Wright, Julie Lounds Taylor, Laura Boteler Humm, Dale E Olsen, Morris D Bell, and Michael F Fleming. 2014. Virtual reality job interview training in adults with autism spectrum disorder. *Journal of autism and developmental disorders* 44, 10 (2014), 2450–2463.

[84] Charles D Spielberger. 1983. State-trait anxiety inventory for adults. (1983).

[85] Jette Sponholz, Eike Born, and Torsten Reiners. 2010. Simulation of Workplace Harassment in 3D Worlds. In *EdMedia+ Innovate Learning.* Association for the Advancement of Computing in Education (AACE), 3862–3865.

[86] Margaret S Stockdale. 1996. *Sexual harassment in the workplace: Perspectives, frontiers, and response strategies.* Sage Publications.

[87] Joachim Taelman, Steven Vandeput, Arthur Spaepen, and Sabine Van Huffel. 2009. Influence of mental stress on heart rate and heart rate variability. In *4th European conference of the international federation for medical and biological engineering.* Springer, 1366–1369.

[88] Katy Tcha-Tokey, Emilie Loup-Escande, Olivier Christmann, and Simon Richir. 2016. A questionnaire to measure the user experience in immersive virtual environments. In *Proceedings of the 2016 virtual reality international conference.* 1–5.

[89] Rebecca A Thacker and Haidee Allerton. 1992. Preventing sexual harassment in the workplace. *Training & Development* 46, 2 (1992), 50–54.

[90] Frank Friedmann Isabelle Viaud-Delmon Thomas Schubert, Holger Regenbrecht. 2016. *Igroup Presence Questionnaire (IPQ) Database.* Retrieved Jan 5, 2021 from http://www.igroup.org/pq/ipq/data.php

[91] Sara Ventura, Georgina Cardenas, Marta Miragall, Giuseppe Riva, and Rosa Baños. 2020. How Does It Feel to Be a Woman Victim of Sexual Harassment? The Effect of 360°-Video-Based Virtual Reality on Empathy and Related Variables. *Cyberpsychology, Behavior, and Social Networking* (2020).

[92] Paul Hendriks Vettehen, Daan Wiltink, Maite Huiskamp, Gabi Schaap, and Paul Ketelaar. 2019. Taking the full view: How viewers respond to 360-degree video news. *Computers in human behavior* 91 (2019), 24–32.

[93] James TR Walters, Jonathan I Bisson, and Jonathan P Shepherd. 2007. Predicting post-traumatic stress disorder: validation of the Trauma Screening Questionnaire in victims of assault. *Psychological medicine* 37, 1 (2007), 143–150.

[94] D. Watson, L. Clark, and A. Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54 6 (1988), 1063–70.

[95] Theresa F Wechsler, Franziska Kümpers, and Andreas Mühlberger. 2019. Inferiority or even superiority of virtual reality exposure therapy in phobias?–A systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard in vivo exposure in agoraphobia, specific phobia, and social phobia. *Frontiers in psychology* 10 (2019), 1758.

[96] Brenda K Wiederhold and Mark D Wiederhold. 2008. Virtual reality for posttraumatic stress disorder and stress inoculation training. *Journal of CyberTherapy & Rehabilitation* 1, 1 (2008), 23–35.

[97] Dennis Patrick Wood, Jennifer A Murphy, Kristy B Center, Carol Russ, Robert N McLay, Dennis Reeves, Jeff Pyne, Russell Shilling, Jack Hagan, and Brenda K Wiederhold. 2008. Combat related post traumatic stress disorder: a multiple case report using virtual reality graded exposure therapy with physiological monitoring. *Studies in health technology and informatics* 132 (2008), 556.

[98] Huizhong Wu. 2017. India says every phone must have a'panic button'by 2017, 2016.

[99] YouGov. [n.d.]. YouGov Omnibus Safety. https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/t8rytgg6ee/Results%20for%20YouGov%20Omnibus%20(Safety)%20072%2020.3.2019.xlsx%20%20[Group].pdf.

[100] Pavel Zahorik and Rick L Jenison. 1998. Presence as being-in-the-world. *Presence* 7, 1 (1998), 78–89.

[101] Chenyan Zhang, Aud Sissel Hoel, Andrew Perkis, and Saman Zadtootaghaj. 2018. How Long is Long Enough to Induce Immersion?. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX).* IEEE, 1–6.