Proceedings Paper:

Promoting Reality Awareness in Virtual Reality through Proxemics

Daniel Medeiros*
University of Glasgow, UK

Rafael dos Anjos†
University College London, UK

Nadia Pantidi‡
Victoria University of Wellington, NZ

Kun Huang‡
Victoria University of Wellington, NZ

Maurício Sousa§
University of Toronto, CA

Craig Anslow‡
Victoria University of Wellington, NZ

Joaquim Jorge¶
Técnico Lisboa, Universidade de Lisboa, PT

ABSTRACT

Head-Mounted Virtual reality (VR) systems provide full-immersive experiences to users and completely isolate them from the outside world, placing them in unsafe situations. Existing research proposed different alert-based solutions to address this. Our work builds on these studies on notification systems for VR environments from a different perspective. We focus on: (i) exploring alert systems to notify VR users about non-immersed bystanders' in socially related, non-critical interaction contexts; (ii) understanding how best to provide awareness of non-immersed bystanders while maintaining presence and immersion within the Virtual Environment(VE). To this end, we developed single and combined alert cues - leveraging proxemics, perception channels, and push/pull approaches and evaluated those via two user studies. Our findings indicate a strong preference towards maintaining immersion and combining audio and visual cues, push and pull notification techniques that evolve dynamically based on proximity.

Keywords: Notifications, Virtual Reality, Human Computer Interaction;Context Awareness, Reality Awareness

Index Terms: Human-centered computing—Empirical Studies in HCI—;——Human-centered computing—Interaction Paradigms—Virtual Reality;

1 INTRODUCTION

Virtual Reality is becoming popular and widespread to the general public, enabling people to be fully immersed in virtual worlds and visualize immersive content. This can be achieved by using the now affordable Head-Mounted Displays (HMDs), which occlude the surrounding environment entirely from visuals and sound, allowing the user to feel transported to a different place. However, being unaware of the physical environment or surrounding people could be harmful to an immersed user [34]. Ideally, one wants to keep the user safe in the virtual environment (VE) without breaking the immersion provided by an HMD.

Previous work has proposed different approaches to protect the user from obstacles (people or objects), such as displaying parts of the real environment [14,20,44], virtual alerts [24,40], or calculating alternative navigation paths [19,31,43] without making users aware of it. "Chaperone" systems of current commercial HMDs, such as the HTC Vive, are now widespread in their role of protecting users from leaving the bounds of the designated interaction space

---

*e-mail: daniel.piresdesamedeiros@glasgow.ac.uk

†e-mail: r.anjos@ucl.ac.uk

‡e-mail: {nadia.pantidi,kun.huang,craig.anslow}@vuw.ac.nz

§e-mail : mauricio.sousa@utoronto.ca

¶e-mail: jorgej@acm.org

where they are safe from obstacles. However, when it comes to non-immersed people sharing the physical space with an immersed user, their relevance is distinct to the one of an obstacle. While obstacles will always be avoided (allowing approaches such as redirected walking to be used [43]), the presence of people in the vicinity of a VR user may have social implications that require a different level of awareness of their existence. Depending on factors such as distance to the user, level of engagement, type of the task being performed in VR, or where the experience is taking place, the user's reaction to the information of "someone else being there" will be different. Thus, other cues may need to be used to allow one to know about it while staying immersed in the VE.

In this work, we propose strategies to enable HMD immersed users to be aware of people in their physical environment without compromising their presence in the virtual space. We tackle this problem by using the concept of proxemics-based interactions and gradual engagement [10, 45], which provides a framework of people's relationship with their space and other people considering different proxemic variables. In our specific scenario, more importance is given to people that are closer to the immersed user and may influence or affects their presence, while users located far away are still made aware but without the same importance as the ones closely located. Many factors may influence people's awareness, including how the VE and the real world are rendered. We define and explore a design space to understand how different types of cues (audio vs. visual, "push" vs. "pull" notifications) influence people's awareness [47]. We then explore these different types of cues to enhance people's awareness in one's space but still maintaining a high sense of presence inside the VE. To accomplish this, we conducted two complementary studies: the first investigated the effects of individual types of alerts on the feeling of presence and subjective metrics regarding usability, awareness, and overall user preferences. The second study, informed by the first study results, evaluated three alert combinations using different perceptual channels (auditory and visual), exploring different levels of intrusiveness in the VE, and a more general versus individualized awareness of non-immersed people. Our results show that users overall preferred using combined alerts due to their complementing capabilities, with an alert that can be checked asynchronously (with "pull" notifications characteristics) to provide overall awareness and a push-based notification to alert bystanders in-close proximity.

2 RELATED WORK

This section addresses relevant research in proxemics, contextual and spatial awareness both in conventional interfaces and VR/AR.

2.1 Proxemics and Gradual Engagement in HCI

Our work considers how people communicate and interact socially. Edward T. Hall [13] introduced the proxemics theory, suggesting that space and distance impact interpersonal relations. Hall refers to proxemics as a silent language since it is a non-verbal and implicit communication device. Proxemics show how people perceive and

use their spatial organization cues to mediate interactions. Furthermore, Hall describes that interpersonal relationships are enacted through proxemic zones (intimate, personal, social, and public interactions). And current research suggests that proxemics help mediate interactions between people and devices for human-to-human or human-to-device interactions [1, 3, 9, 10, 17, 29]. Marquardt et al. [28] generalized cross-device information transfers by proposing the gradual engagement design pattern. The suggested pattern formulates how devices can gradually engage users by revealing different information as a function of proximity and orientation. That is, the user interfaces progressively move through various stages affording gradual engagement degrees. Previous work has used proxemics to make people and technology more visible through light, animation, and recognizable shapes in software applications [1, 46, 50]. For example, Vermeulen et al. [45] introduced Proxemic Flow, a method using floor light indicators to guide people interacting with public floor displays, and hence, using proxemic-related information to indicate the state of the interaction. On a similar note, for portraying people in mixed reality spaces, Sousa et al. [42] presented "social bubbles" to convey presence and provide awareness of proxemic interactions between remote people through floor projections. Previous research also suggests that the way people interact with each other and their virtual surroundings in VEs match the proxemics fundamentals [6,7]. Sanz et al. [37] formalized the concept of virtual proxemics when studying how people avoid obstacles and are aware of other people in VEs. The authors show that people's interactions with proximal virtual objects are comparable to real-life behaviors. We can conclude that proximity-based cues are a natural communication device to provide feedback about the environment and the surrounding people.

2.2 Context-aware notifications in Mixed Reality

Human attention is limited [2, 27], so the use of alerts to notify people needs to be carefully designed to both alert and direct people's attention to situations or objects present in both the interface and/or the outside world [39]. Notifications are commonly sub-divided into two categories: *push* and *pull* [5]. *Push* notifications are the ones that notify the user even without their consent and *pull*, the ones that the user can choose when to be notified. Previous work investigated push and pull type notifications and their implications for a range of devices and applications [11, 49]. Audio cues are common forms of push notifications and are used to alert people about urgency situations [12], where sounds with higher volume and frequency are more effective in alerting people. Auditory notifications are also effective in communicating urgency in driving contexts, which demands high levels of attention from drivers [30]. These types of notifications have also been successfully used to communicate the presence of objects in space when in VR [52].

Visual cues are also an effective form of alerting people and redirecting attention to objects on the interface [15]. These objects are often used as push notifications, which are overlaid on the interface to notify promptly about specific objects [25]. Sheikh et al. [38] use arrows overlaid into 360 videos to redirect people to the vital part of the content being displayed at a given time. Visual cues are also shown to be the most effective in locating objects in VR scenes when compared to both auditory and vibrotactile cues [4].

2.3 Enabling spatial awareness in Mixed Reality

More recently, researchers have started to consider ways to enable people to be aware of their surroundings while immersed in Virtual Reality. These works primarily focus on the surrounding environment and ways to alert the immersed user of obstacles present in the physical room. Classic examples include the HTC Vive, which shows a stylized rendering of the physical room. This rendering can be rather disruptive and only alerts people when they are in the borders of a controlled space. However, a controlled environment is not always available, and VR may be used in locations such as people's living room or in public places.

Previous work also addressed the potential of providing awareness of physical objects in the room in both VR and AR. An example of such a solution is the Reality Check [14], which presents a space that merges virtual and physical environments, enabling people to be aware of their surroundings while allowing for correct rendering of virtual objects and moving elements. A follow-up study [21] also suggested that proxemics-based interactions enable awareness of obstacles in the physical setup both on AR and VR. Other approaches also include the use of redirection of users' virtual path so that users do not collide with both static objects [19, 22, 40, 41] and moving people [43]. Von Willich et al. [48] on the other hand, investigate the effects of rendering fidelity of tracked bystanders, which shows that a 3D abstract avatar is not statistically less accurate than a point-cloud representation, while being less disruptive.

Aside from the need for awareness to avoid obstacles and ensure safety, the ability to be aware of people in the physical space may have social implications and influence the way people interact inside the VR space. For instance, when non-immersed bystanders move closer to VR-users, it may indicate a potential desire to interact with them or even intentionally disrupt the VR experience. When non-immersed people are far-away, they may co-inhabit the space with no intention to interact. Still, their presence may need to be acknowledged by the immersed users in particular as they transition to areas closer to the immersed user. Wolfe et al. [51] presented a first attempt to describe the use of proxemics for interactions for non-immersed users in VR. However, their work limited proxemics' usage to a highly controlled environment and immersed users had no awareness of people outside the VR space. Follow-up work also enable the use of visual and audio cues [33, 35, 36] and multimodal notifications including visual, audio and haptic feedback [8, 26] to alert VR-users of bystanders. Due to the human's visual dominance, visual and audio cues still seem to be the most efficient way to notify VR users of bystanders [8]. Ghosh et al. [8] go a step further and investigate the use of notifications with multiple channels and its interruptions' effects for notifying people in VR. However, they focused on the general effects of such notifications and proxemics interactions for people in the same space were not covered properly. As shown in the previous work, the delivery of alerts in Mixed Reality environments has been predominantly of the push type notification. This is likely due to the emphasis on developing ways to alert for danger to ensure immersed users' safety. Still, it leaves unexplored the space of pull notifications, which may allow for more agency and immersion to be preserved in non-safety critical environments. Our work aims to explore this gap through the investigation of both push and pull type notifications. In this paper, we further explore the effects of proxemics-based alerts to make VR-users aware of non-VR immersed users without completely disrupting their level of presence within the VR space.

3 OUR APPROACH: ISOLATED CUES

As presented in the previous section, different visual and audio cues have been proposed for VR/AR scenarios, showing a different impact on the users' presence. Also, depending on the goal/task proposed, different cues have seen more success than others. For our first exploratory study, we designed various cues that capture the type of elements proposed by previous academic work and traditional game metaphors. We applied the proxemics framework to each cue's design, changing their intensity or representation according to the position of the non-immersed person. This section presents each of the individual cues, how to build them and how the proxemics framework has been applied to their design.
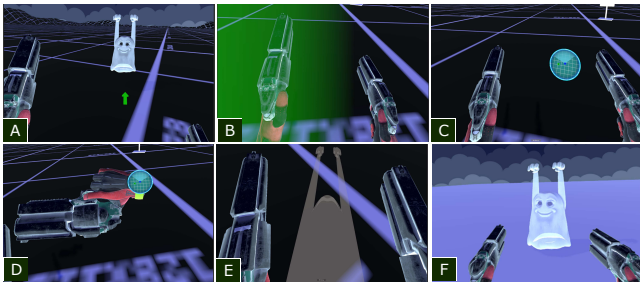
Figure 1: Visual alerts tested; Overlays : (A) Arrow3D (B) Color-Glow (C) Minimap Overlay; World Elements: (D) Watch minimap (E) Shadow (F) Ghost avatar

3.1 Overlays: Push notifications

Similarly to previous work in both mobile and conventional devices [15, 23, 25], we implemented visual notifications that are displayed directly to the user's field of view. We consider these "push notifications" [5, 47], as the user is made aware of non-immersed bystanders without his request. Each tested notification varies not only visually, but also on the granularity of the displayed information. We implemented three alternatives (see Figure 1), with increasing precision regarding bystanders' positions.

Color glow: A glow on the edge of the HMD which indicates that there is a person in that direction (left, right, above for behind, and below for people locate in the front of the user). This notification can appear in one or multiple sides of the HMD screen, depending on bystanders' positions. This notification can appear in one or multiple sides of the HMD screen, depending on bystanders' positions. The glow is a 3D semi-transparent rectangular object. We adjusted the position and the size of it based on the user's perspective, which takes 18% of the HMD screen size. The color of the glow indicates what proxemics region the closest person on that side is; we use green for public space, yellow for social space, and red for personal space (e.g. two people to the right, one in the personal space and one in the public space, the glow on that side will indicate red for personal space). This is the least granular alert, as it simply alerts of the presence of one or more people on a certain direction.

3D arrow: A set of 3D arrows displayed at a fixed distance from the VR user's head. Each arrow points to an individual bystander. The color of the arrow indicates which proxemics region the non-immersed person is. This cue allows the immersed user to know the the direction where a person is, but not their exact location or movements.

Mini-map: A 2D Mini-map displayed at a fixed distance from the VR user's head, in the middle of its field of view containing the position of all bystanders relatively to the VR user, as in a video-game. Bystanders are visualised as dots on the mini-map and color-coded according to proxemics zones. We included a grid of one square meter to aid users to better estimate their physical distance to bystanders.

3.2 World elements: Pull notifications

Alternatively, we implemented alerts that need to be prompted by the user, which we would classify as "pull notifications" [5, 47]. These require some sort of action to be triggered by the user in the virtual environment for them to be made aware of the presence of non-immersed people. We used easily identifiable objects in the scene, as suggested by Ghosh et al. [8] in order to reduce visual search, while matching the visual style of elements to increase presence and not to disrupt the task being performed. In our implementation, they are represented by world elements that are increasingly visible to the user(Figure 1) so to explore a range on the 'pull' continuum. The watch is the most obvious pull notification, the shadow less so but

still a pull one as due to the shooting task implementation, users have to purposefully look down to be able to see it, and the ghost is on the border between pull and push.

Watch: a Mini-map that is attached to the user's wrist, which is the least intrusive alert, only visible when prompted by the user. When checked, it contains all the positions for all non-immersed users. It also contains a color-coded map of bystanders according to the defined proxemics zones.

Ghost: A ghost avatar at the exact position of each non immersed person. The ghost avatar is visible when it is in the personal space or when the VR user's looking directly to it in the social space.However, it will be nearly transparent if it looked at other directions, which means the bystander is not trying to interact with the VR user. It is the clearest reference to the specific position of bystanders, and also the most intrusive of the three cues.

Shadow: A shadow on the floor starting at the position of the non-immersed user, and stretched towards the VR user. The bigger the shadow is, the closer the other person is. The shadow was implemented using a flat texture and was resized according to the bystander's position. The user will only see this alert if he/she looks at the floor. This is a middle ground between the watch and the ghost alert, as it is easier to prompt than the Mini-map, but less visible than the ghost.

3.3 Audio cues: Continuous vs discrete

In addition to the visual cues, audio cues have been proposed by previous work as a way to get the users' attention in tasks that are mainly visual [40, 52]. We were interested in evaluating the users' perception of continuous vs discrete notifications, and if spatial sound was accurate enough to allow users to locate non-immersed people. While spatial audio allows one to roughly locate sound sources while in VR, we wanted to evaluate how accurately it can communicate the position of a non-immersed person when dissociated from any visual representation. Thus, for audio cues we implemented two alternatives.

3DBeep: Constant 3D beeping sound, spatially located at each bystander's position. Pitch and frequency were increased discretely at the different proxemic areas, sounding more alarming when in the personal space(0.6 s interval), as opposed to more casual when in the public space (2s interval).

Footsteps: 3D Sound of footsteps, activated according to movement from each non-immersed bystanders' positions. In this alert, we adjusted the volume of the footsteps according to the bystander's proxemics regions using a logarithmic roll-off curve.

4 FIRST USER STUDY

The aim of this first user study was to evaluate the implemented cues based on presence, user preference and awareness of bystanders while immersed in VR. Unique to our study is the full immersion of the HMD-based VR system (different from [1]) which does not allow people to rely on real world cues. This user study also focus on awareness of real people and its potential social implications.

Instead of using real people as bystanders, we recorded eight one minute movement scripts with one to two bystanders.These were recorded using HTC Vive Trackers and were recorded in the same room where the user study was conducted. By using recorded movement scripts we were able to ensure that all the participants experienced the same interactions intentions evenly and guarantee safety due to COVID-19 restrictions in place. The movement scripts mimicked real-world interactions, from a passive interaction such as just standing far-away from the immersed user to a more active scenario, where bystanders came closer to the immersed user in an attempt to interact with them. Given that there was no interaction between immersed and pre-recorded non-immersed users, the recorded scripts represent realistic real world scenarios, allowing our findings

to be generalized. These scripts were delivered together with the alert cues in a randomized manner to avoid learning effects.

Differently from previous work that focused on non-demanding tasks [35], we designed a visually demanding task, with both audio and visual effects that require enhanced attention from users. Our task consisted of a shooting game, where users were asked to shoot moving targets (disks) thrown at them, similar to [24]. This task was subdivided into three sub-tasks, which differed in how the moving targets were presented at the users, making people look in different directions. In Sub-task one, cannons shot disks from the behind towards the users, forcing them to look forward; in sub-task two, objects were thrown up, forcing people to look up; and the final sub-task had disks thrown up and around the users. All tasks were performed with game sound effects and music to counterbalance cognitive effects for both sound and visual-based alerts. We used the HTC Vive Pro as the visualization device and used the controllers to enable interaction within the VE. Users were asked to remain in a fixed position to guarantee that the alerts would work as intended.

4.1 Methodology

We recruited 12 participants with ages varying from 20 to 54 years old. Most of the users had previous experience with VR technology. First, participants were presented with the consent form and then a brief explanation about the study's overall purpose, the task, and the individual alerts and were allowed to ask any questions. After that, participants were asked to position themselves in the center of the room and gear up with the HMD and controllers. When ready, participants started the shooting task, which was performed eight times (one per alert). Each task took one minute, with a pause in-between, where users verbally answered a short questionnaire (see below). The questionnaire gathered their experience with the said alert. The order of alerts was counter-balanced to reduce learning effects. After that, users filled a profile questionnaire that included demographic information and questions regarding their previous experience with VR and AR technology. Finally, we conducted a semi-structured interview to capture participants' subjective responses regarding the alerts, tasks, and overall experience.

4.2 Questionnaire Results

We collected subjective metrics through questionnaires. The questionnaires comprised seven questions, where one of those asked the perceived number of simulated people in the environment. The further six were 5-Likert scale questions, which were asked after completing the task with one of the alerts: (1) I felt like I was actually there in the VE (presence) (2) I noticed the alerts (3) the alert was effective to indicate people in the room (4) the alert was effective in indicating proximity (5) The alerts interrupted my focus from the task (6) I like the alert. The results for these questions are summarized in Figure 3. For the questionnaire metrics, as they are a discrete variable, we used a Friedman non-parametric test with Wilcoxon Signed-Ranks test posthoc tests.

4.2.1 Overlays

Regarding overlay alerts, users felt that these types of alerts had a similar effect in eliciting a high sense of presence (q2), being noticed by the users (q3), detecting people in the room (q4) and interrupting the focus from the shooting task (q5). When asked about which one they liked the most users preferred overall the 3DArrow alert when compared both with the Colour Glow (Z=2.154 p=0.031) and Minimap overlay alert (Z=-2.846 p=0.004).

Two of these alerts were amongst the most noticeable alerts, with the exception of the Minimap alert. Both the Color Glow and the Arrow3D were more noticed than both sound alerts (Beep: Z=-2.345 p=0.019; Footsteps:Z=-2.354 p=0.019), both of them were also more noticed than the watch alert (Colorglow: Z=-2.558 p=0.011;
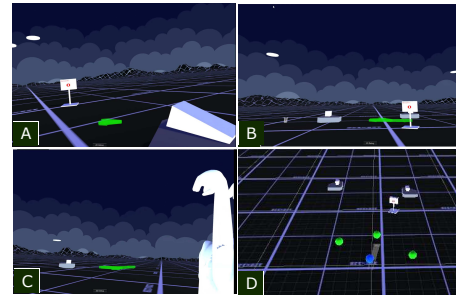

Figure 2: Shooting tasks: (A) Cannons shooting towards the user (B) Cannons shooting up (C) Cannons shooting up and around the user (D) Overview of the virtual environment; the green spheres denote recorded bystanders, and the color indicates its proximity to the immersed user, which is represented by a blue sphere.

Arrow3D:Z=-2.326 p=0.02). The Colorglow was also more noticed when compared to the shadow alert (Z=-2.326 p=-2.124).

People also found the Color glow and arrow3d more effective for indicating people in the room when compared to the 3D Beep (Arrow3d: Z=-2.09 p=0.037; Colorglow: Z=2.404 p=0.016). The Colorglow was also found to be more effective than both ghost (Z=-2.121 p=0.034) and shadow (Z=-1.976 p=0.048) world elements.

Finally, participants liked these type of alerts except for the Minimap that had the lowest score of them. When compared to other types of alerts the Arrow3D was more liked the 3D Beep (Z=-2.459 p=0.014).

4.2.2 World elements

The shadow was the least preferred of the tested alerts in this category (Ghost: Z=-2.459 p=0.014; Watch: Z=-2.348 p=0.02).The ghost alert was more noticeable (q3) when compared to the Watch Minimap (Z=-2.326 p=0.02), and was the most effective on identifying people in the room when compared to both watch (Z=-2.555 p=0.011) and shadow (Z=-2.457 p=0.014).

When comparing to other types of alerts, the ghost was seen as the most effective overall by participants. It was more noticed (q3) than the Beep (Z=-2.326 p=0.02), more effective to identify people in the room (q4) than both the 3D Beep (Z=-2.934 p=0.003) and the Minimap overlay (Z=-1.956 p=0.049). People also liked this alert more than the beep (Z=-2.539 p=0.011). In a direct comparison between Minimap alerts tested, participants preferred the Watch Minimap in comparison with the Minimap overlay (Z=-2.588 p=0.011) as the latter was said to be disruptive and difficult to focus.

4.2.3 3D Audio

There was no significant differences between the two spatial audio alerts tested. These type of alerts were the least noticed by the users as compared to overlays and world elements alerts. In this respect, the Beep performed the worst, being significantly less noticed than both the Arrow3d (Z=-2.086 p=0.037) and colorglow (Z=-2.345 p=0.019). The footsteps alert was also less noticed than the Color glow alert (Z=-2.354 p=0.019).

4.3 Accuracy

In addition to the likert scale questions, we also asked to users how many non-immersed users they perceived to be in the room, as part of the questionnaire. We applied a Friedman non-parametric test to test for main effects followed by a Wilcoxon Signed ranks test as the results were not normally distributed.

We consider accuracy as the percentage of right answers. We found that the Arrow3D was the notification that was the most effective on notifying people. The arrow3d was found to be statistically better than the 3DBeep (Z=-2.0700 p=0.038), ColorGlow
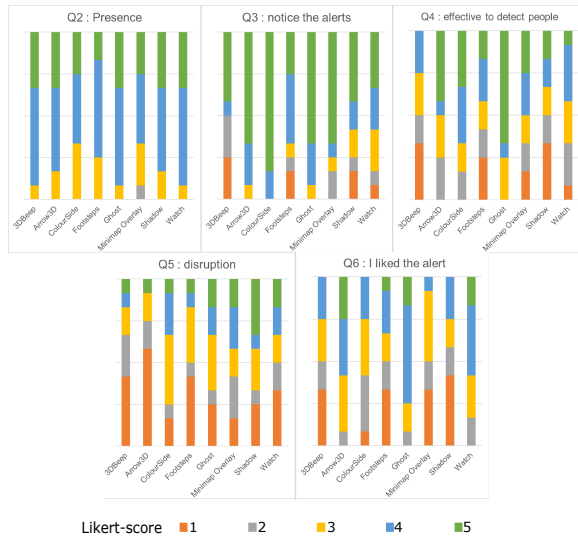
Figure 3: User preferences grouped by question for the First Study

(Z=-2.271 p=0.023), Footsteps (Z=-2.887 p=0.004), Ghost Avatar (Z=-2.000 p=0.046), Shadow (Z=-2.236 p=0.025), and Watch (Z=-2.000 p=0.046). Additionally, the footsteps alert was statistically worse than the Minimap (Z=-2.121 p=0.034) and shadow (Z=-2.236 p=0.025).

When ranking the alerts in terms of accuracy we noticed that the Arrow3D was the most accurate (100% of participants), followed by the minimap (with 75%, or 9 participants) followed by the Watch and the Ghost avatar (66.6% or 8 participants). This were followed by shadow and 3DBeep with 7 correct answers, Colorglow (with 50% of correct answers). The least accurate was found to be the footsteps alert, with only 3 participants (25%).

With the Color glow, we noticed a tendency of overestimating the number of perceived bystanders, the same behaviour was also noticed by the Minimap and shadow. On the other hand, some of the alerts caused an underestimation of the number of perceived bystanders such as the Minimap, shadow, 3DBeep and footsteps. By analyzing the Figure 4, we can see that both audio alerts seemed to cause a significant underestimation in comparison to the others.

4.4 Interview Findings

In this section, we report our data analysis findings from the participants' interviews from study 1. More specifically, we provide participants' synthesized accounts on the overall experience per each alert and their insights on the value of having alert cues in immersive scenarios and which circumstances and features are considered more
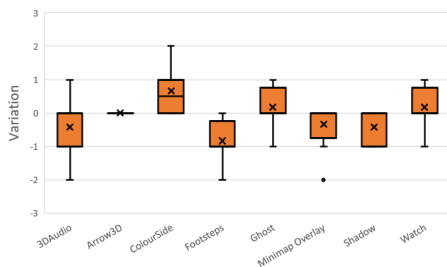

Figure 4: Accuracy on perceived bystanders on the first study: In this graph we differentiate underestimation from overestimation of bystanders

critical or appropriate for those.

4.4.1 Preferred alerts

When asked which single alert was their preferred overall, all our alerts got at least one mention with the exception of the 3d beep. Interestingly the majority of them discussed how they see more value in having a combination of two or sometimes three alerts that complement each other and provide notifications towards different aspects such as number of people, proximity and actual location. When citing combined versions of alerts, the 3D beep was mentioned. *"(. . .) The ghost is good because you can see it, the beep is useful to indicate movement (...)" "The arrow shows you where someone is when you can't see them directly (P1.12)"*

"(...)I would have a Minimap to know where people then I'd start to introduce the beep but if they are really really close then you would the visual element to see where they are so know I need to start to be careful" (P1.8)

4.4.2 Evaluating individual alerts

What was evident across our analysis (interview and questionnaires) of participants' reporting of individual alerts, is that there is no 'one size fits all' and in this respect all type of alerts received positive and negative remarks across a range of aspects.

Overlays: Of the three overlay type alerts the 3D arrow was the most well received across all participants, which confirms the quantitative results of being well perceived and accurate. *"(...) arrow is the most favourite, always showing. I know there are people, where the people are." (P1.5)*

Despite the overall positive feedback for the 3D arrow, some small issues were reported by a few participants, indicating this alert could be complemented: *(...) "I wasn't sure if the length of the arrow was indicating anything so I felt, it was actually quite hard to read, it was like reading a clock a little bit." (P1.3)*

The side colour was found familiar to people who regularly play shooting games and effective in terms of providing a general awareness of whether people were in the room and whether they were in proximity. Improvements included reducing its size (P1.12), replicating it in both eyes (P1.11) and showing the glow only when non-immersed users were in their personal space (P1.9).

Participants found the mini map a useful and 'familiar or easy to learn' alert system and reported it as less intrusive for VR tasks. However, the current implementation of the mini map, as part of a "heads up display", made participants 'squint' to see it in detail, which caused an underestimation effect as pointed by the accuracy results. This alert was found to not be ideal and the majority of the participants suggested it needed to be changed.As mentioned in the quantitative results, the watch implementation was overall preferred.

World elements: The watch was overall well received, even though it has lower accuracy than the Minimap. Several participants pointed out that it was well suited for the VR environment as it doesn't break the immersion and equally fitting to the shooting game task context.

Still, a few participants found that having to look at the watch disrupted the focus from the shooting task and some found it somewhat difficult to quickly decipher the information on the map due to orientation of the watch and the small scale. This again alluded to providing more than one way to make the user aware of non-immersed bystanders.

"Yes, I think at least the idea of having a mini map is, because the mini map on the wrist watch is good but it unnecessarily interrupts whatever you are doing. So it is good to be able to check that but it can't be the only thing you are relying I think." (P1.12)

The ghost alert sparked strong reactions both during the time participants were carrying out the shooting task and later in the interview. During the study we observed participants laughing,trying to avoid the ghost and being surprised or spooked by it. Their

comments reflected those reactions, as they highlighted liking how effective the ghost was as an alert but equally the level of disruption it incurred in terms of their immersion and focus in the shooting task, but it gave them a good reference of size and scale. *"The ghost is good because you can see it (...) it is the only thing that actually gave me a sense 'yes it is there' and not somewhere between me and infinity. The problem I guess is field of view because if the ghost is behind you, you won't notice it(P1.12)*

The shadow alerts were found to not be as noticeable compared to the other world element alerts. This to some extent was attributed to the context of the shooting task as most of the targets participants had to shoot were located in the upper field view of the world and to 'see' the shadows users had to look down to the ground.

"(...) even though shadow didn't disturb my sight but just i had to look down to see the shadow" (P1.6) "With the shadow you have to remember if the long shadow means close or far away." (P1.2)

Sound Alerts: The sound alerts were found by the majority of our study participants to be not as distinct and noticeable due to different reasons. The beep was found to be clashing with the audio track of the shooting game, and hard to discern specific positions.

"The only issue with the audio is that the beep is kinda similar to the noises of the game. It would be nicer if it was less similar but that is why it is also good to have several cues." (P1.12)

While the footsteps was found to be little descriptive of closeness. *"I like the footsteps better than beep but in terms of footsteps, i couldn't tell that the people were coming to me or not, just i could be aware where people are. But I didn't know how close they are" (P1.6)*

As a result, participants scored these alerts fairly low (Figure 3) and their accuracy in the experimental conditions involving the audio alerts was also low as we can see in Figure 4. The footsteps were slightly better received compared to the 3d beep whose sound was found more intruding.

Despite the aforementioned issues, the majority of participants still saw significant potential in having sound cues together with visual cues as mentioned previously, as they had their value in providing awareness. *"[Best for general awareness] the beep. I was immediately aware that someone was there because it is just an annoying sound. Especially as it resembles alarms so I am "that is bad, I got a be aware of.""(P10)*

4.4.3 Valuing alerts but prioritising the immersive experience

While all participants appreciated the value of having some form of alert while in VR, it was evident that priority was given to the immersive experience itself. The majority of our participants stressed that it was more important to them to have an, as uninterrupted as possible, immersive experience, which they identified as the main reason for doing a task in VR in the first place.

"(...) when I put the VR on it is to experience a VR world, and having a constant stream of updates of what is happening in the real world can throw you out a little. It does kinda contradict some of the points of VR" (P1.2)

As a result all the identified benefits of benefits of having alert notifications were weighed against the level of disruption these might cause. *"Like it was, if you took the screen glow and made it flash, it would take away from the experience. But I think for me, an ideal setup would be someone enters the room, you hear the notification, and then it is up to you if you want to check the watch and to see where that person is (...) but I think beyond that, it starts to interfere with the experience." (P1.2)*

4.4.4 Push vs Pull

Another aspect that participants highlighted as they debated alert features was the notion of control; with some participants wanting alert notifications to be 'pushed' on them through the VR environment and others wanting to be able to 'pull' those as/when they wished while having an ambient, calm awareness. These differing

preferences mirror interaction practices and styles that people hold across or are familiar with from other digital devices and applications (e.g. games) that they interface with or from their everyday life experiences.

"Yes [it is better to be notified without having to check] because otherwise it is distracting all the time because you have to do two things."(P1.4)

"I dunno if it is from other games or VR tasks, but it is just I get to do it with my hands, I get to control it with my hands, and I like that aspect of it. (...) yes although I do recognise that that of course has problems with if you forget to check you are not aware of your surroundings (...) I would still prefer the watch because I can move it to where it is comfortable for me at the current moment." (P1.10)

5 SECOND USER STUDY: REVISED AND COMBINED CUES

The results of the first study gave us insight on individual alert cues and how they can be used. However, we felt there was more to be explored in terms of how the intrusiveness of the alerts fared across, and furthermore, how the push-pull notifications can be used in combination. The majority of participants expressed that some of the alerts complemented each other during the post-test interviews which suggests that shortcomings of certain alerts could be well covered by providing a different cue at the same time. These could be visual and auditory combinations, but also push and pull combinations, where the push notification alerts you of something generic, that can be further clarified by the user actively pulling that information or vice-versa. We can summarize the findings of each of the alerts as follows:

3D arrow: Always visible and accurate on direction, less so on distance.

Side Color: Very inaccurate for number of people, viable for indicating general direction of a "threat".

Minimap: Accurate, but VR Heads up displays (HUD) force users to focus on information that is close to the eyes in depth, making it difficult for people to focus on the interface.

Watch: Perceived as the better alternative to the minimap. The smaller accuracy occurred due to the need of actively checking the watch, while remaining aware of the task itself.

Shadow: Less noticeable as being on the floor. Lower accuracy was due to a similar reason as the watch. Unclear on distance/direction when out of FOV.

Ghost: Accurate when in field of view, although most intrusive and disturbing of main task. Informs exact position.

3D Beep: Not accurate on direction, but accurate on distance.

Footsteps: More accurate on position than on distance.

Given this, for our second study we proposed three different combinations of alerts which were found to compliment each other's shortcomings, while also being in a spectrum between non-intrusive and intrusive, allowing us to explore how having two alert cues instead of a single cue affects the feeling of immersion in the VR experience.

Combination 1 (Beep + Watch): As the beep (audio push) was successful in notifying distances through audio, we believe it is a good combination with the watch (visual pull), which when checked, provides accurate information about positions. This allows users to choose when to check the watch, according to the distance and number of beeps. We consider this to be the least disturbing combination, as no visual elements are presented unless purposefully checked, and the beep was noted as being drowned out by the in-game effects at times (mainly when faint and far away) which works in favor of making a less intrusive alert.

Combination 2 (Footsteps + Shadow) : Footsteps (audio push) were successful in indicating direction of a non-immersed person, while the shadow could indicate position clearly when in the field of view. For this second study we considered the shadow as being a

Visual Pull notification, as the shooting task required mainly looking up, thus looking at the shadow needs to be a purposeful action.

Combination 3 (Ghost + Arrow3D) : This combination corresponded to the most intrusive combination of alerts in our user study, but also pointed to be the most accurate by the results of the first user study. In this combination, we use the Ghost avatar, which shows the exact 3D position of the bystander non-immersed user. Alone, this notification was not as effective as users need to actively search for it; Because of that, we combined this alert with the 3D Arrow, which solves this problem by pointing to the bystander position, even if they are located behind the HMD-wearer.

For this study we recorded different movement scripts, recorded from real people movements and included two to four people.

5.1 Methodology

The methodology of the second user study was similar to the first one. In this study we recruited 12 subjects, who had not participated in the previous study. In this user study, participants performed the shooting task three times (each lasting 1') for each of the alert combinations, with randomized order of movement scripts and alert combinations to avoid learning effects. For this study we recorded different movement scripts, recorded from real people movements and included two to four people. At the end of each alert combination trial, participants were asked verbally to rate different aspects tested. After testing all combinations, users completed a profile questionnaire and were then asked to rank the three combinations in terms of their overall preference, effectiveness to provide general awareness, effectiveness to indicate proximity and level of disruption from the IVE, followed by a semi-structured interview.

In the next sections, we report on both the quantitative and qualitative results resulting from our analysis. We subdivided the quantitative analysis in form of questionnaires and accuracy metrics, and then we present the subjective findings from the interview separately.

5.2 Questionnaires

For the second user study we used a similar questionnaire to the first study, but modified to reflect on the combinations of alerts. This was comprised of 6 5-Likert scale questions, which were asked verbally to users so they do not need to take off the headset in-between conditions. These questions were the following : (1) how many people did you notice? (2) i felt like i was actually there in the VE (presence) (3) I noticed the alerts (3) the alert was effective to indicate people in the room (4) the alert was effective in indicating proximity (5) The alerts interrupted my focus from the task (6) I like the combination of alerts. As mentioned earlier, we also included a set of ranking questions after participants completed all experimental conditions and before the semi-structured interview.

Regarding user preferences, we did not find statistical significance in any of the asked questions. Summarized results gathered from the questionnaires are summarized in Figure 5. We performed an additional Two-one sided t-test to test for equivalence and found equivalent results in terms of presence for all tested conditions (C1-C2: p=0.006; C1-C3: p=0.009;C2-C3 : p=0.009). We also noticed equivalent results on noticing the alerts between C1 (Watch+3DBeep) and C3 (Ghost+3DArrow) (p=0.009). When asked directly to rank these aspects in the post-study questionnaires, we found that the C2(shadow + footsteps) was the least favourite of the three, with nine of the 12 participants ranking them as their 3rd least preferred. This was also said to be the least effective in notifying about other people in the room (9 out of 12 ranking it as third place), the least preferred to indicate proximity (8 out of 12 ranked third place), but it was not seem as disruptive as the C3 (Ghost+3DArrow), with 7 of the users ranking it as their second most disruptive alert.

Regarding the other two combinations, results show that the C1 (Beep+Watch) was the overall preferred, followed by the C3 (3DArrow+Avatar). When asked about which was the most effective in
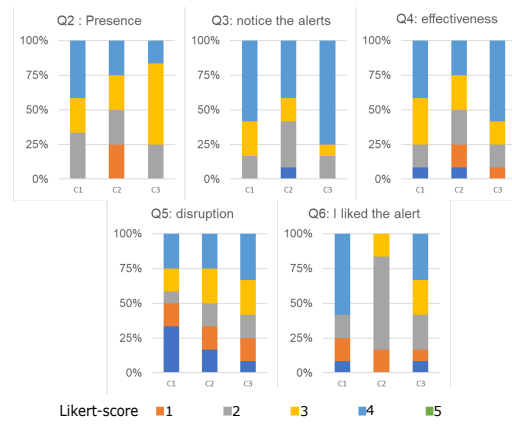

Figure 5: Questionnaires Results for the 2nd study split by question. C1 is the combination with the 3D Beep and Watch, C2 - Shadows and Footsteps and C3, 3D Arrow and Ghost Avatar

notifying about other people, participants preferred the Combination 3 (8 out of 12), followed by combination 1 (Beep+Watch) and lastly the combination 2. The Combination 1 was the preferred in terms of indicating proximity with a slight advantage (6 out of 12) over the Combination 3 (4 out of 12). These preferences are summarized in Figure 6.

5.3 Accuracy

For accuracy we considered both absolute and relative accuracy of the correct number of perceived bystanders. We noticed statistical significance between the Combination 3 (Ghost+3DArrow) (Median=0 IQR=0) and Combination 2 (Shadow+Footsteps) (Median=-1 IQR=1)(Z=-2.000 p=0.046). No statistical significance was found between the other combinations regarding accuracy regarding perceived bystanders.

5.4 Interview findings

Similar to the first study, short semi-structured interviews were conducted with each participant after they completed the experimental part of the study. In these interviews participants elaborated on their reasoning for the answers they gave in the questionnaires and provided valuable information regarding their overall experience, their assessment of the alerts and their perceived value. Interviews were audio recorded, transcribed and analysed using content analysis. Below, we present key outputs from the interview data analysis which provided significant and useful information towards the design of VR alert notifications, and clarified or deepened the results from the quantitative analysis.
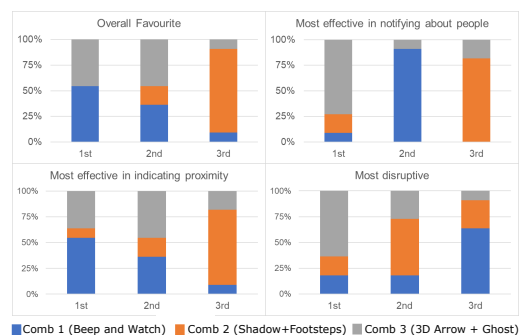

Figure 6: Results from the ranking questionnaire of the second study.

Overall, having a combination of alerts was better received compared to having a singular one; and the majority of participants expressed strong preference towards having an alert combination of visual and audio cues. While this may not be immediately evident from the quantitative data, in their interviews, participants showcased significant preference for combination 1 that combined audio and visual cues against 3 that only had visual cues. The reasoning provided was that such combination is more effective as i) audio and visual cues complement each other and can serve different notification purposes (e.g. one for proximity, one for number of people) and ii) it accommodates individual preferences and styles in terms of how people register and process informational stimuli e.g. some people identified themselves having strong visual preference or mentioned not registering audio cues when they are in "the flow of a focused activity".

"having an auditory cue makes you aware that something is happening and you can't always make sense of it, so having the visual cue really helps in terms of 'oh I can count how many people there are'.They help each other in a way..." (P10)

Similar to the first study, participants stressed the value of having an alert system in place and mentioned specific application areas for that, such as gaming at home/in a public space, but also activities inspired by the recent pandemic context such as working and teaching from home using VR. Preserving immersion was still a high priority especially for gaming and entertainment applications. This further explains participants' aforementioned preference towards comb. 1 - despite finding comb. 3 as the most accurate - as combination 3 was perceived as particularly disruptive to immersion and task due to the ghost avatar. At the same time, having effective ways to break that immersion were also discussed with varied audio and visual combos depending on the type of activity or the urgency of the alert.

"if they are far away it is good to have the arrows or a model of where they are in the room. But when they come closer to your personal space, then the audio the beeping would be a very good option to have" (P2)

As seen from the above, combinations of audio and visual cues were found particularly suitable to address these varied circumstances and in this respect, further features and representations of alerts were brought forward that provided useful directions for future research and development. For example, participants discussed about alerts that were more 'natural' (such as using the footsteps to indicate someone walked in the room while the other person is in a VR meeting) but bear the tradeoff of being harder to interpret and requiring more time to learn. Participants explained how having to spend more time to 'learn' what the alert is conveying was adding to the cognitive load of the activity and was not preferable. This also explains why participants,as mentioned earlier, preferred Combination 1 more than 3 (and even less 2).

"And then I found the footsteps quite hard to distinguish, you could maybe hear two sets of footsteps but I wasn't sure if it was one (person) or more" (P4)

In addition, participants considered and debated on the merit of having different representations of the visual and audio cues depending on the context of use. For example, several participants suggested having universal, standardised alerts across all applications; clustering visual information in public space situations, when there is a big number of people and/or when they are not in close proximity:

"better to find some kind of visualisation that can show aggregates of things,you have colour already showing proximity so something else maybe a model of clusters of people or how big is the group" (P10)

6 CONCLUSION

HMDs for VEs provide full-immersive experiences to wearers. However, completely isolating people from the outside world places them in unsafe situations. Existing research proposed some alert-based solutions to address this issue. Our research builds on these endeavors

on notification systems for VR environments but explores a different approach. We are interested first in devising alert systems for notifying VR users about non-immersed bystanders' presence for socially related, non-critical contexts of interaction. And second, understanding how best to provide awareness of bystanders while keeping a high presence and immersion within the VE. We developed different alert cues to address these, both isolated and combined - leveraging proxemics, perception channels, and push/pull notification strategies - that we evaluated in two user studies.

Findings from the first study indicate that while participants acknowledged the value of alerts, there is no 'one size fits all' technique – all registered negative and positive remarks over various aspects. Most importantly, though, participants' responses indicated the complementary value of having a combination of 2-3 alerts using different cues (audio and visual) and with both push and pull together, where we could have a less disturbing alert as a first notification, and a clearer one when a non-immersed person was closer and vice versa. The first study's findings informed the second study's design, where we combined alert cues and evaluated those using a similar experimental design. In the second study, we found that participants preferred having a combination of audio and visual cues, as indicated by the qualitative results. Combined alert cues can support different attention modes and allow for a better distribution of cognitive load in a visually demanding task, which is common in VR and also was the case in our study. Accuracy-wise, explicit visual cues performed slightly better, as expected. However, qualitative analysis of the interviews and users' preferences indicated that less disruptive combinations (combo 1: audio beep with a wrist mini-map in our case) were preferred by most users, despite not being significantly worse than the more explicit alerts (Combo 3).

In both studies, participants were keen on not breaking immersion during a VR experience. Undisturbed immersion becomes even more critical in cases where there are no imminent danger or safety issues. In these situations, participants strongly preferred staying immersed and having control and agency over when to break VR to check on bystanders. Other work in Ubicomp and HCI reported a similar strong preference by people for controlling when to check their devices and receiving notifications on demand [16, 18, 32]. Still, to our knowledge, this has not been found before in a VR context.

While we can argue that this approach is fairly general, attention-getting mechanisms are contingent on the immersed users loci of attention and therefore inherently task-dependent. We conjecture that there might be an expanded design space relating attention getting devices to intrinsic of the task at hand. This generalization will be the subject of future studies. Furthermore, our analysis provides a nuanced understanding of what are 'natural' alert cues. Participants discussed cues that are familiar from other digital contexts (e.g., gaming), real-world activities (e.g., looking at your watch or hear footsteps as someone walks in) or just 'universally' easy to interpret (colored arrow), which cut down on the learning effort and do not overburden the cognitive load of the VR task at hand.

Finally, our findings point to directions for future research regarding alerts' features and representations. Which include investigating different cues such as vibro-tactile, haptics and smell; and incorporating identity information or detection as part of the alerts, while still keeping disruption low. Another promising avenue is looking for ways to cluster visual information for many people, such as in public space situations. Indeed, this research becomes imperative as VR and AR applications and devices become ever more pervasive in everyday life.

REFERENCES

[1] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, pp. 121–130, 2010.

[2] D. E. Broadbent. A mechanical model for human attention and immediate memory. *Psychological review*, 64(3):205, 1957.

[3] F. Brudy, C. Holz, R. Rädle, C.-J. Wu, S. Houben, C. N. Klokmose, and N. Marquardt. Cross-device taxonomy: survey, opportunities and challenges of interactions spanning across multiple devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–28, 2019.

[4] T. Chen, Y.-S. Wu, and K. Zhu. Investigating different modalities of directional cues for multi-task visual-searching scenario in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–5, 2018.

[5] K. Cheverst, K. Mitchell, and N. Davies. Exploring context-aware information push. 6(4):276–281, Jan. 2002. doi: 10.1007/s007790200028

[6] B. R. Fajen and W. H. Warren. Behavioral dynamics of steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2):343, 2003.

[7] M. Gérin-Lajoie, C. L. Richards, and B. J. McFadyen. The negotiation of stationary and moving obstructions during walking: anticipatory locomotor adaptations and preservation of personal space. *Motor control*, 9(3):242–269, 2005.

[8] S. Ghosh, L. Winston, N. Panchal, P. Kimura-Thollander, J. Hotnog, D. Cheong, G. Reyes, and G. D. Abowd. Notifivr: Exploring interruptions and notifications in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1447–1456, 2018. doi: 10.1109/TVCG.2018.2793698

[9] S. Greenberg, N. Marquardt, T. Ballendat, R. Diaz-Marino, and M. Wang. Proxemic interactions: the new ubicomp? *interactions*, 18(1):42–50, 2011.

[10] J. E. Grønbæk, M. S. Knudsen, K. O'Hara, P. G. Krogh, J. Vermeulen, and M. G. Petersen. Proxemics beyond proximity: Designing for flexible social interaction through cross-device interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–14. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376379

[11] X. Guan and B.-Y. Choi. Push or pull? toward optimal content delivery using cloud storage. *Journal of Network and Computer Applications*, 40:234–243, 2014.

[12] E. C. Haas and J. Edworthy. Designing urgency into auditory warnings using pitch, speed and loudness. *Computing & Control Engineering Journal*, 7(4):193–198, 1996.

[13] E. T. Hall. *The hidden dimension*, vol. 609. Garden City, NY: Doubleday, 1966.

[14] J. Hartmann, C. Holz, E. Ofek, and A. D. Wilson. Realitycheck: Blending virtual environments with situated physical reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.

[15] C. Hoareau. *Elaboration et évaluation de recommandations ergonomiques pour le guidage de l'apprenant en EVAH: application à l'apprentissage de procédure dans le domaine biomédical*. PhD thesis, Brest, 2016.

[16] E. Horvitz, J. Apacible, and M. Subramani. Balancing awareness and interruption: Investigation of notification deferral policies. In *International Conference on User Modeling*, pp. 433–437. Springer, 2005.

[17] S. Houben, J. Vermeulen, C. Klokmose, J. Schöning, N. Marquardt, and H. Reiterer. Cross-surface: Challenges and opportunities of spatial and proxemic interaction. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, pp. 509–512, 2016.

[18] S. T. Iqbal and E. Horvitz. Notifications and awareness: a field study of alert usage and preferences. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pp. 27–30, 2010.

[19] K. Kanamori, N. Sakata, Y. Hijikata, K. Harada, K. Kiyokawa, et al. Walking assist method for vr zombie. In *2019 12th Asia Pacific Workshop on Mixed and Augmented Reality (APMAR)*, pp. 1–7. IEEE, 2019.

[20] K. Kanamori, N. Sakata, T. Tominaga, Y. Hijikata, K. Harada, and K. Kiyokawa. Obstacle avoidance method in real space for virtual reality immersion. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 80–89, 2018. doi: 10.1109/ISMAR.2018.00033

[21] H. Kang and J. Han. Safexr: alerting walking persons to obstacles in mobile xr environments. *The Visual Computer*, pp. 1–13, 2020.

[22] H. Kang, G. Lee, and J. Han. Obstacle detection and alert system for smartphone ar users. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359996.3364256

[23] H. Kang, G. Lee, and J. Han. Obstacle detection and alert system for smartphone ar users. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359996.3364256

[24] J. Lacoche, N. Pallamin, T. Boggini, and J. Royan. Collaborators awareness for user cohabitation in co-located collaborative virtual environments. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–9, 2017.

[25] Y.-C. Lin, Y.-J. Chang, H.-N. Hu, H.-T. Cheng, C.-W. Huang, and M. Sun. Tell me where to look: Investigating ways for assisting focus in 360 video. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2535–2545, 2017.

[26] C. Mai, M. Hassib, and C. George. Like elephants do: Sensing bystanders during hmd usage. In *Proceedings of the CHI 2017 Workshop on Amplification and Augmentation of Human Perception, May 07, 2017, Denver, CO, USA. Copyright is held by the owner/author (s)*, 2017.

[27] F. Maringelli, J. McCarthy, A. Steed, M. Slater, and C. Umilta. Shifting visuo-spatial attention in a virtual three-dimensional space. *Cognitive Brain Research*, 10(3):317–322, 2001.

[28] N. Marquardt, T. Ballendat, S. Boring, S. Greenberg, and K. Hinckley. Gradual engagement: facilitating information exchange between digital devices as a function of proximity. In *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, pp. 31–40, 2012.

[29] N. Marquardt, K. Hinckley, and S. Greenberg. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pp. 13–22, 2012.

[30] D. C. Marshall, J. D. Lee, and P. A. Austria. Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human factors*, 49(1):145–157, 2007.

[31] S. Marwecki, M. Brehm, L. Wagner, L.-P. Cheng, F. Mueller, and P. Baudisch. Virtualspace-overloading physical space with multiple virtual reality users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–10, 2018.

[32] A. Mashhadi, A. Mathur, and F. Kawsar. The myth of subtle notifications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pp. 111–114, 2014.

[33] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A dose of reality: Overcoming usability challenges in vr head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2143–2152, 2015.

[34] M. McGill, J. Williamson, A. Ng, F. Pollick, and S. Brewster. Challenges in passenger use of mixed reality headsets in cars and other transportation. *Virtual Reality*, pp. 1–21, 2019.

[35] J. O'Hagan and J. R. Williamson. Reality aware vr headsets. In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays*, pp. 9–17, 2020.

[36] J. O'Hagan, J. R. Williamson, and M. Khamis. Bystander interruption of vr users. In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays*, PerDis '20, p. 19–27. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3393712.3395339

[37] F. A. Sanz, A.-H. Olivier, G. Bruder, J. Pettré, and A. Lécuyer. Virtual

proxemics: Locomotion in the presence of obstacles in large immersive projection environments. In *2015 IEEE Virtual Reality (VR)*, pp. 75–80. IEEE, 2015.

[38] A. Sheikh, A. Brown, Z. Watson, and M. Evans. Directing attention in 360-degree video. 2016.

[39] R. Soret, P. Charras, C. Hurter, and V. Peysakhovich. Attentional orienting in virtual reality using endogenous and exogenous cues in auditory and visual modalities. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, pp. 1–8, 2019.

[40] M. Sousa, D. Mendes, and J. Jorge. Safe walking in vr. In *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359997.3365737

[41] M. Sousa, D. Mendes, and J. Jorge. Safe walking in vr using augmented virtuality. arXiv:1911.13032 [cs.HC], 2019.

[42] M. Sousa, D. Mendes, D. Medeiros, A. Ferreira, J. M. Pereira, and J. Jorge. Remote proxemics. In *Collaboration meets interactive spaces*, pp. 47–73. Springer, 2016.

[43] Q. Sun, A. Patney, L.-Y. Wei, O. Shapira, J. Lu, P. Asente, S. Zhu, M. Mcguire, D. Luebke, and A. Kaufman. Towards virtual reality infinite walking: Dynamic saccadic redirection. *ACM Trans. Graph.*, 37(4), July 2018. doi: 10.1145/3197517.3201294

[44] I. Valentini, G. Ballestin, C. Bassano, F. Solari, and M. Chessa. Improving obstacle awareness to enhance interaction in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 44–52. IEEE, 2020.

[45] J. Vermeulen, K. Luyten, K. Coninx, N. Marquardt, and J. Bird. Proxemic flow: Dynamic peripheral floor visualizations for revealing and mediating large surface interactions. In *IFIP Conference on Human-Computer Interaction*, pp. 264–281. Springer, 2015.

[46] D. Vogel and R. Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pp. 137–146, 2004.

[47] V. von Jan, S. Bertel, and E. Hornecker. Information push and pull in tactile pedestrian navigation support. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, MobileHCI '18, p. 55–62. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/ 3236112.3236120

[48] J. von Willich, M. Funk, F. Müller, K. Marky, J. Riemann, and M. Mühlhäuser. You invaded my tracking space! using augmented virtuality for spotting passersby in room-scale virtual reality. In *Proceedings of the 2019 on Designing Interactive Systems Conference*, DIS '19, p. 487–496. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3322276.3322334

[49] D. Weber, A. Voit, P. Kratzer, and N. Henze. In-situ investigation of notifications in multi-device environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 1259–1264, 2016.

[50] A. D. Wilson and H. Benko. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, pp. 273–282, 2010.

[51] M. Wölfel, D. Hepperle, A. Siess, and J. Deuchler. Staging location-based virtual reality to improve immersive experiences. *EAI Endorsed Transactions on Creative Technologies*, 6(21), 2020.

[52] J. Yang, P. Sasikumar, H. Bai, A. Barde, G. Sörös, and M. Billinghurst. The effects of spatial auditory and visual cues on mixed reality remote collaboration. *Journal on Multimodal User Interfaces*, pp. 1–16, 2020.