

Towards Balancing VR Immersion and Bystander Awareness

YOSHIKI KUDO, Tohoku University, Japan

ANTHONY TANG, University of Toronto, Canada

KAZUYUKI FUJITA, Tohoku University, Japan

ISAMU ENDO, Tohoku University, Japan

KAZUKI TAKASHIMA, Tohoku University, Japan

YOSHIFUMI KITAMURA, Tohoku University, Japan

Head-mounted displays (HMDs) increase immersion into virtual worlds. The problem is that this limits headset users' awareness of bystanders: headset users cannot attend to bystanders' presence and activities. We call this the HMD boundary. We explore how to make the HMD boundary permeable by comparing different ways of providing informal awareness cues to the headset user about bystanders. We adapted and implemented three visualization techniques (Avatar View, Radar and Presence++) that share bystanders' location and orientation with headset users. We conducted a hybrid user and simulation study with three different types of VR content (high, medium, low interactivity) with twenty participants to compare how these visualization techniques allow people to maintain an awareness of bystanders, and how they affect immersion (compared to a baseline condition). Our study reveals that a see-through avatar representation of bystanders was effective, but led to slightly reduced immersion in the VR content. Based on our findings, we discuss how future awareness visualization techniques can be designed to mitigate the reduction of immersion for the headset user.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)** → **Interaction paradigms**

KEYWORDS: HMD Boundary, Virtual Reality, Informal awareness

ACM Reference format:

Yoshiki Kudo, Anthony Tang, Kazuyuki Fujita, Isamu Endo, Kazuki Takashima, and Yoshifumi Kitamura. 2021. Towards Balancing VR Immersion and Bystander Awareness. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 5, ISS, Article 484 (November 2021), 22 pages, <https://doi.org/10.1145/3486950>

1 INTRODUCTION

People wear head-mounted displays (HMDs), or headsets, to experience virtual reality content. Because the virtual content fills the headset user's visual field and auditory experience, the real world disappears, affording rich immersive experiences due to the removal of distractions. Yet headset users physically remain in spaces that may contain other people and objects (e.g. [36, 46]).

This work is supported in part by JSPS Grants KAKENHI (19KK0258).

Author's addresses: Yoshiki Kudo, Tohoku University, Japan, ykudo@riec.tohoku.ac.jp; Anthony Tang, University of Toronto, Canada, tonytang@utoronto.ca; Kazuyuki Fujita, Tohoku University, Japan, k-fujita@riec.tohoku.ac.jp; Isamu Endo, Tohoku University, Japan, isamu.endo.t8@alumni.tohoku.ac.jp; Kazuki Takashima, Tohoku University, Japan, takashima@riec.tohoku.ac.jp; Yoshifumi Kitamura, Tohoku University, Japan, kitamura@riec.tohoku.ac.jp.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 License.

© 2021 Copyright held by the owner/author(s).
2573-0142/2021/11-ART484. <https://doi.org/10.1145/3486950>

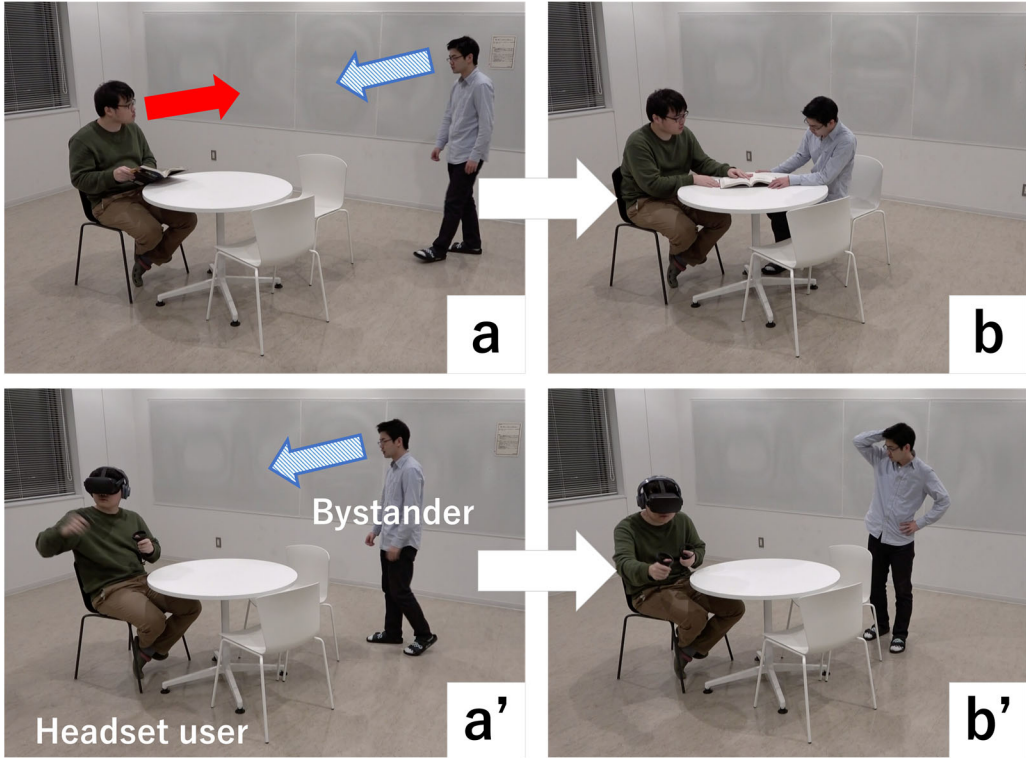


Fig. 1. The HMD Boundary: In (a'), the headset user is oblivious to the co-present bystander; (b') thus the bystander has no awareness of what the headset user is doing. Entering into conversation is awkward.

This introduces a problem: headset users are blind to what is happening in their physical context, which can make it cumbersome or dangerous to engage with nearby bystanders [17, 21, 29, 34].

In contrast, in everyday life, we subconsciously collect information about the people and environment around us, which we use to build informal awareness [13, 42] loosely defined as “knowing who is around and available for interaction” [16]. For instance, we hear the sound of a door opening, or footsteps of approaching people; we notice people as they come into our field of view, as well as what they are doing. This informal awareness of others’ presence and proximity—whether they are around, what they are looking at—informs us what they are doing, who they are interacting with, whether they are busy or interruptible, and so forth (e.g. [10, 13, 24]). We use these continuous signals to maintain an understanding of the world around us: when others slightly shift their shoulders toward us, this may imply interest in engagement—though if we notice that this is combined with turning toward someone else in the environment, we understand this movement is intended to engage with another person. Thus, we use these continuous signals that make up informal awareness to mutually and gradually negotiate whether and how we engage in further interactions with one another.

The problem is that these signals are not readily available between headset users and bystanders, because the virtual world is disconnected from the physical world [12, 14, 15, 29]. Once a person dons the HMD, their awareness of the location and activities of people who enter or move around their physical space disappears. The HMD thus imposes a barrier—one we call the ‘HMD boundary’ (Figure 1). This boundary creates unsafe conditions when headset users and bystanders occupy shared physical spaces (e.g. [11, 33, 34, 36]), and creates confusion for the headset user when the social environment changes outside of the VR simulation [23].

In this paper, we explore how to make the HMD boundary permeable by providing informal awareness cues to the headset user about bystanders in the environment. Prior work in CSCW

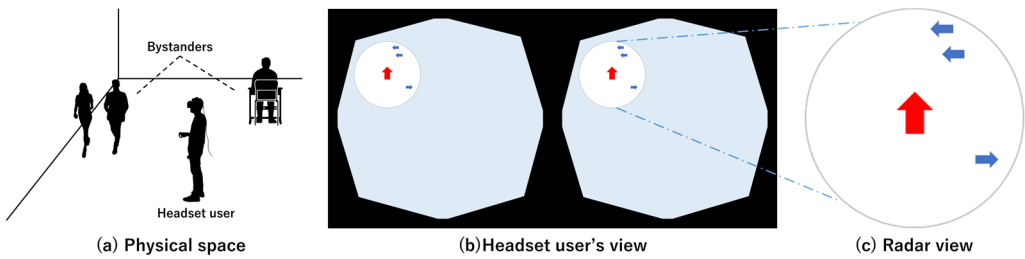


Fig. 2. This schematic illustrates how our approach works, where: (a) the headset user may be sharing the physical space with other bystanders; (b) the headset user's view of the virtual world includes an awareness interface widget; (c) a detail view of the Radar widget shows how the bystanders position and orientation (in blue) relate to the headset user's (in red).

studying awareness have identified ways of sharing presence and location of others effectively through visual cues in 2D workspaces (e.g. [16]). Drawing on this work, we adapted three such awareness visualizations for the HMD context (Avatar, Radar and Presence++). Each of these vary in two main ways:

1. Where is the awareness information presented? (situated in the visual scene, or separate from it)
2. To what extent is the awareness information abstracted? (literal information, or symbolic)

Each of these techniques has been demonstrated to be effective in its own right (e.g. [16, 43]), though it is unclear how these apply to the context of visualizing bystanders a headset user may choose to ignore, and it is unclear what factors designers should consider when choosing between them. Our intention was to identify visualizations that allow the headset user to be immersed in the VR content while being spatially and socially aware of their physical surroundings. While Rzayev et al. conducted a similar exploration for digital notifications [37], we focused on the cues represent bystanders around the headset user.

In our prototype designs, we continuously monitor the space around the headset user, capturing the presence, location, and orientation of bystanders. As illustrated in Figure 2, this is revealed to the headset user with cues designed to provide “just enough information” so they can continuously mediate their interactions with bystanders in the physical environment.

While others have begun tackling this design problem (e.g. [14, 15, 29]), our approach differs from prior work in two fundamental ways: first, while others focus on explicit communication across the HMD boundary (e.g. [14, 15]), our interest is on providing *awareness* of bystanders; second, we explicitly *compare* different continuous indicators of bystanders' presence in relation to immersion. For instance, whereas Willich et al. [43] show that Avatar-style visualizations are effective to promote awareness, this visualization only works for a narrow field of view and may negatively impact immersion. Other techniques that place the visualization on the visual periphery (e.g. Radar View [16]) may help the headset user maintain immersion since the awareness information is no longer in the focal visual area. By comparing different visualization techniques, our work aims to find techniques that balance the tension between allowing the headset user to remain immersed in the virtual content while maintaining an awareness of her surroundings [29]. Finally, our work assumes the near-future scenario where the headsets are not visually or auditorily “leaky”: that is, no external noise enters the headset user's experience (as has been explored by others [34]).

We conducted a controlled comparative simulation study to identify effective visualizations to enable informal awareness with twenty headset users. In this study, participants were immersed in headset VR experiences while simulated bystanders passed by. These simulated bystanders variously stopped with the expectation of interacting with the headset user or not (e.g. the bystanders might simply be passing through, or stopped in the environment to talk with someone

else, and so forth). We compared different variations of our approach, where headset users could “see” these bystanders using different visualization variations. Through the evaluation, we assessed whether participants could ascertain the presence of bystanders, and their “interest” in initiating interaction. Our results show that it is possible to design cues to allow headset users to stay aware of the bystanders, and some cues are more disruptive to the immersive experience than others.

We make four contributions. First, we relate the HMD boundary problem to classic problems of awareness, allowing researchers to bring this rich vocabulary to bear on this new problem. Second, we adapt existing visualization techniques for enabling informal awareness across the HMD boundary for headset users. Third, we develop a new simulation-based study approach for evaluating techniques for supporting informal awareness of bystanders. Fourth, we provide results of a study that compare different design factors (situated vs. separate; literal vs. symbolic) may be used in visual cues to provide awareness to VR users. Our work can aid designers looking to convey this continuous awareness information about bystanders to VR users in the visual domain, to augment existing work exploring the aural domain (e.g. [33]).

2 RELATED WORK

Two areas of related research set the stage for our work: prototypes to enable interaction across the HMD boundary, and research exploring how informal awareness enables people to gracefully engage and disengage with one another.

2.1 Interaction across the HMD Boundary

The principal challenge presented by the HMD boundary is that it separates headset users from bystanders who are outside of the virtual environment [29]. Headset users are unaware when bystanders are present or interested in engaging in interaction (e.g. [12, 29, 33]). Bystanders are unable to understand what headset users are experiencing in the VR environment (e.g. [14, 19, 35]). Current commercial implementations (e.g. Steam VR’s “Knock Knock”) provide a crude binary mechanism where a bystander needs to explicitly notify the headset user of her engagement by pressing a button on the keyboard, which then interrupts the headset user with a discrete UI notification. Yet, based on prior literature, we know that awareness and engagement is far more fluid and requires support that allows for ambiguities, and subtlety [16].

Many researchers have explored supporting communication across the HMD boundary via viewports, supporting explicit interaction, or by stylizing the view of the other world. McGill et al. provide a canonical example of the viewport approach, where the headset user is supplied a view of the outside world by blending a semi-transparent video capture of the physical world into the VR space [29]. Hartmann et al. proposed RealityCheck which enables headset user to directly see physical world by integrating the physical world inside a virtual world [17]. Similarly, for bystanders, virtual content can be projected on physical walls [19, 20], floor [41], on a separate smartphone [35], or a tablet [40]. These viewports can enable explicit communication through drawing and annotation [15], or in an augmented space [14].

Another approach provides cues to the headset user that limits or stylizes real-world information. Budhiraja et al. render part of a webcam view of the physical world (e.g., a moving hand) [2]. He et al. implemented five different visualization cues (indexical, symbolic, and iconic with three emotions) based on two dimensions (vividness and emotion) [18]. Mai et al. proposed concept of subtle cues using vibration, electricity and pressure, though they did not implement and evaluate the systems [26]. Kang et al. proposed a safety assistance system “SafeXR” which alerts walking VR headset users to obstacles through visualization cue [21]. Ghosh et al. explore a number of ways to present notifications through various audio, visual and haptic cues [12]. Zenner et al. stylize these notifications to fit within the virtual world scene (e.g., email delivered by a virtual avatar) [47]. Rzaev et al. explored four different visual notifications in three VR

Category	Element	Question
Who	Presence	Is someone here?
	Identity	Who is it?
What	Action	What are they doing?
Where	Location	Where are they?
	Gaze	What are they looking at?

Table 1. To address informal awareness across the HMD boundary, cues should help answer these questions for the headset user (adapted from [16]).

user an understanding of the social environment—people that are present, whom they are interact with, and the extent to which they may be interested in engaging with the headset user. O’Hagan et al. [34] show that headset users can maintain some of this type of continuous awareness in the using open-ear headsets; however, in our work, we assume the use of closed-ear headsets for a fully immersive experience. The closest work in this space is Willich et al., where they design and compare three different representations of physical passersby in the virtual environment [43], including a video-based representation, an avatar-based representation, and a point cloud style representation. In our approach, we explore abstracting this information about bystanders, where the presentation of the information is separated from the actual location of the bystanders. Prior work has suggested that such symbolic presentations can be interpreted more rapidly [16]. Similarly, Medeiros et al. explored push and pull notification systems by combining audio and visual cues. By introducing proxemics theory, they changed the presentation of cues based on distance. Our cues are not designed based on proxemics zones, and instead are simply push and pull notifications based on different visual characteristics [30].

2.2 Informal Awareness between Distant Collaborators

Informal awareness has long been of interest to researchers exploring casual interaction between distance-separated collaborators. The goal was to design systems that allowed collaborators to maintain subtle, unspoken awareness of others’ presence, availability and activities. This awareness in turn afforded lightweight casual interactions across various media channels. For instance, somewhat similar to the HMD boundary viewport work described above, the design of always-on video media spaces broke the distance boundary by providing an interactive audio/video viewport linking the separate physical spaces of collaborators [7, 8, 10]. Collaborators kept aware of what others were doing through the link, and used that information to decide when to initiate communication through this link.

Gutwin and Greenberg [16] distill the components of informal awareness by framing them as a set of questions that a user might want to know about others in the environment (Table 1). How these questions are answered depends primarily on context. For instance, most messaging systems rely on very simple awareness cues that indicate presence or activity. WhatsApp [48], Facebook Messenger [49] and similar systems typically show which contacts are online via presence indicators (e.g. a green circle). These indicate whether someone is around to be interacted with. Similarly, these messaging apps can show whether a contact is currently typing a message. This visual activity indicator provides awareness and helps mediate expectations of immediacy in the interaction. Similar to physical distance, the HMD boundary—unless somehow made permeable—impairs how people gather awareness of one another. The research above shows how using low-level awareness mechanics transmitted across the boundary can mitigate the boundary of physical distance.

As described by Gutwin & Greenberg [16], these visualizations of awareness information can vary in how they are presented and visualized along two dimensions, placement and presentation.

environments [37]. Yet, these alerts, notifications are discrete and jarring in the virtual environment, far from the nuanced ways in which people subtly engage one another through their proxemic relationships [4, 22].

Whereas this prior work focuses on supporting explicit communication across the HMD boundary, we focused on designing subtle informal awareness cues that give the headset

Placement concerns whether the visualization is situated or separate, where a situated cue is visualized where the information originates, whereas a separate cue is visualized somewhere distinct from where the information is created. For instance, Google Docs provides situated cues for collaborators' mouse cursors (they represent where the remote collaborator is working), but separate cues for collaborators' viewpoints (where these are visualized on a thumbnail of the entire document). Presentation concerns whether the visualization is literal or symbolic: a literal cue is one that is shown in the same way the data is gathered, whereas a symbolic cue is more abstract. Willich et al. [43] display information about bystanders using a literal presentation style. The video avatars do not abstract the bystanders' appearance; a symbolic presentation, for instance, might focus only on the presence or absence of bystanders (e.g. like online presence indicators of instant messaging tools). Further, the placement style of these visualizations was strictly situated—that is, the information about the bystanders is visualized where it originates from; in contrast, a separated placement allows for placing this information somewhere else in the interface. This separated placement where placed within the field of view may increase awareness efficiently [37]. Thus, we go beyond Willich et al. [43] by showing how other corners of the design space from [16] are also effective in improving awareness, while maintaining the immersive aspects of the VR experience.

3 BALANCING INFORMAL AWARENESS AND IMMERSION

In current systems, the VR experiences are intended to fully immerse the headset user. The idea is to design them such that she can focus on the content without disruption, and can feel immersed and present in the experience. This is expressed in standard tools to measure VR experiences (e.g. [38, 39]), where items ask the user about her experience of the content vs. her experience of the real world. Such measures explicitly set up “well-designed” immersive experiences to be those that block out the real world such that the headset user feels present solely in the virtual environment.

The starting point for our work is the increasingly common scenario of a headset user engaged in a VR experience, where she physically shares the space with others in the environment (e.g. [34, 36]). As we can see from the following scenario, each bystander has unique intentions and motivations, particularly as they relate to their intentions toward the headset user: only one of them is interested in interacting with the headset user.

Hannah the headset user moves her arms around and walks around as she engages in her immersive first-person shooter experience, dodging enemies, aiming, and shooting at them with her gun. The room is a busy place. John stands nearby to talk with Kaori. Ken walks by to get a book on the bookshelf; because he cannot see the tracking space, he assumes he walks clear of it (but he does not). Luis enters the room with the intention of interrupting Hannah to talk to her.

In this canonical scenario, the onus is solely on bystanders to initiate and negotiate the interaction with the headset user; our interest is in returning some of the control over this para-interaction back to the headset user without overly disrupting the VR experience. As others have shown (e.g. [2, 12, 29]), it is possible to provide headset users with a view to the outside or to notify the headset user when bystanders are present and interested in interaction. Yet, these approaches may unnecessarily be disruptive to the immersive experience. Furthermore, notification mechanisms are crude and discrete, and in contrast to the graceful ways we negotiate our understanding of others' presence and interest in the physical world [4, 22]. How can we design mechanisms to provide the headset user with sufficient awareness of bystanders without unnecessarily disrupting the immersion? And, how should this approach be mediated by the nature of the interactive VR content—for instance, does a quieter, calm VR experience without too much interactivity demand a different approach than a loud, fast and highly interactive VR experience (e.g. a first-person shooter game)?

To address this design tension, our approach focuses on providing informal awareness of the bystanders through cues presented to the headset user. We begin with elements that make up informal awareness: who is present (presence), where are they (location), and what are they doing (orientation, trajectory)? Our goal is to understand how to present cues to the headset user that answer these questions in a way that they can absorb this information without disrupting the immersive experience.

4 DESIGN VARIATIONS FOR INFORMAL AWARENESS AND IMPLEMENTATION

We adapt and re-purpose existing visualization designs that present informal awareness information to the headset user about bystanders (Figure 3). The three variations are drawn from video games and prior work (e.g. [16, 43]). We chose these visualizations because they present informal awareness information (i.e. presence, location, orientation and trajectory of bystanders) in different ways (i.e. literal vs. symbolic), and places this information differently within the view (i.e. situated vs. separate). The Avatar view provides a virtual “see through” into the physical space (a literal, situated presentation), the Radar widget provides a top-down view of the space (a literal, separated presentation), and the Presence++ widget provides an iconic view of the bystanders (a symbolic, separated presentation).

4.1 Design Variables

Whereas Gutwin & Greenberg’s placement and presentation framework of visual cues [16] were developed for shared 2D visual workspaces, we re-interpret those ideas for 3D environments where users are wearing VR headset displays.

With regard to placement of cues:

Situated: In situated condition, we show the cues at original position. In this case, we show the cues to represent bystanders where they are in relation to the headset user’s position and orientation. This is close to real world situation, allowing the user to easily understand how far away the bystander is from the headset user, and to understand bystander’s engagement to the headset user. However, if the headset user does not face to the bystander, the cues will be displayed out of headset user’s field of view (the headset user would miss the bystander).

Separated: In separated condition, we show the cues far from original position, but show the cues within the headset user’s field of view like heads-up display. The headset user can see the cues any time. So, as argued by [37], the headset user easily notice the cues compared to situated ones. But, the cues provide information far from original position. So, it may difficult to interpret the bystander’s intentions. In addition, heads-up display may cause disruptiveness, since it appears within the field of view if it is not necessary.

With regard to presentation of cues:

Literal: In literal condition, we present the information about the bystander in the same way it is captured. In this case, we show location and orientation of the bystander continuously. Thus, the headset user can infer trajectory information based on the continuous signal, as well as other subtle cues based on orientation and trajectory. This mimics real life.

Symbolic: In symbolic condition, we extract specific information from the original tracking data, and make inferences based on it. In this case, we abstractly provide location and orientation information of the bystander to the headset user. For instance, we can provide the location and orientation in ordinal or binary steps (i.e. close or far, and facing to headset user or not.) This may be easier to interpret compared to literal cues; however, the amount of the information is lower.

4.2 Prototype Design

We implemented three different visualization cues by referring the design variables. Each cue was implemented as an example. While we have several candidates to design the cues (e.g. design cues based on VR contents [18]), we simply implemented the visualization cues.

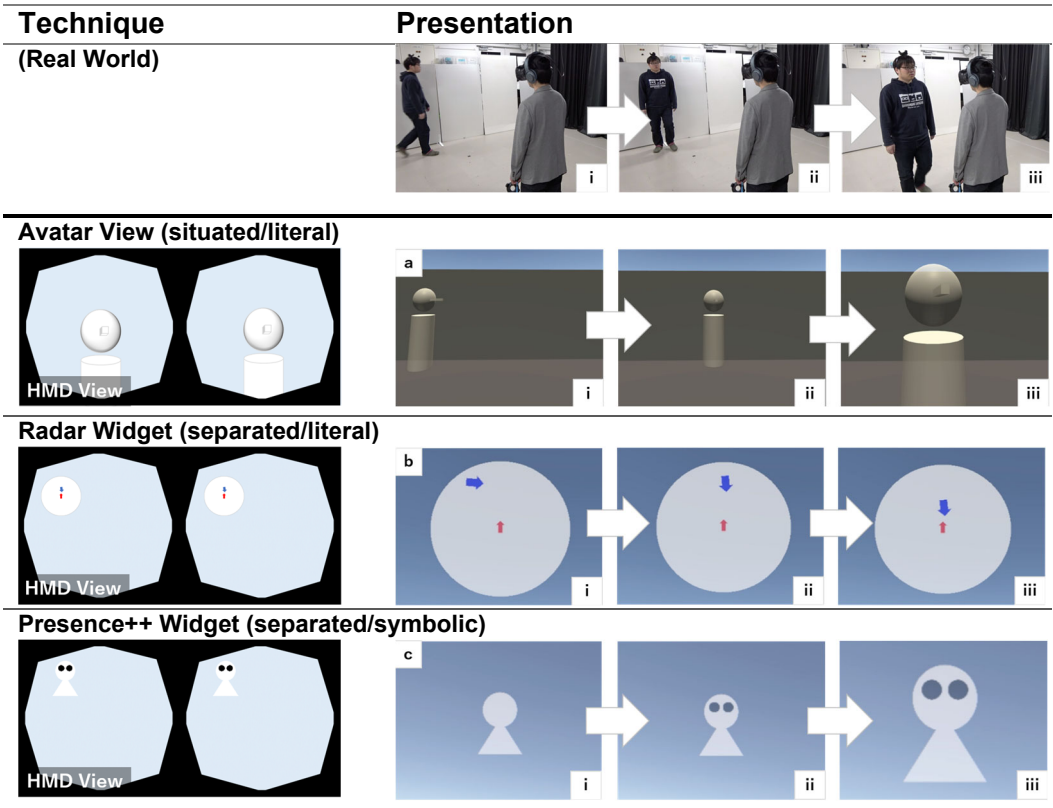


Fig. 3. Each of three design variations shows what happens: (i) in the presence of a bystander, (ii) when the bystander faces the headset user, and (iii) when the bystander is near the headset user. Note that Radar and Presence++ are zoomed, detail views; they appear much smaller.

Avatar-View (situated and literal). In this variation, we re-implement the design proposed by [43], where bystanders are represented by avatars in the virtual environment (Figure 3a). The avatars are cylinders with a spherical head and nose), and their position in the virtual environment maps 1:1 to the bystander’s location in the physical space, and the avatar’s nose orientation maps 1:1 to the bystander’s head orientation. When the bystander is sitting, their avatar is shorter.

Rationale: We designed the avatar view as a type of comparison case to prior literature [33, 43], where our design focuses on showing presence, location, and orientation—basic awareness elements from [16]. This approach provides a high level of accuracy and is more familiar to users in the sense that it is similar to how real-world informal awareness cues are gathered. However, the avatars may occupy a large part of the virtual environment (and the headset users’ field of view), and this may be overly disruptive—particularly in situations where there are a lot of bystanders around, and especially when these bystanders do not want to engage with the headset user. Further, one challenge with this visualization is that headset users are only aware of bystanders who are within the narrow “field of view”.

Radar Widget (separated and literal). In this variation, the system presents a “radar view”, a top-down view of the physical space that shows the presence and orientation of bystanders with blue arrows (Figure 3b). The radar is centered around the headset user’s location and head orientation, and the bystander arrows are visualized relative to this. Thus, as the headset user steps forward or turns to the right, a bystander in the radar widget would appear to move and circle around the headset user.

Rationale: This design is reminiscent of the radar view from prior research studying awareness [16], and is familiar to many gamers, since the interface widget is used in many video games. These games use the radar as an “overview” of the local environment, where the user can see the presence, location, and orientation of friends/foes within a game display. Since the headset user can see the Radar widget anytime, this may help maintain awareness; however, the HUD-style may disrupt the VR experience since the information is provided continuously.

Presence++ Widget (separated and symbolic). This design is fully symbolic, we show core elements of the informal awareness through icons (Figure 3c). The icon has two sizes: when a bystander enters a tracked space, the icon is small; when the bystander enters the play area, the icon becomes large. The eyes of the icon are mapped to the relative orientation of the bystander to the headset user—an icon that has eyes means the bystander is looking toward the headset user (when the bystander is looking within 45° of the headset user); an icon without eyes means the bystander is looking away from the headset user. We show up to three icons—one for each of the three closest bystanders.

Rationale: This variation focuses on the core elements of informal awareness, and visualizes this information abstractly. Given the goal is to avoid distracting the headset user from the immersive content, we present this information in the least obtrusive way possible. While the information presented here is less rich, it may be sufficient for many situations, allowing users to read the icons in a timely way.

In our implementation, we track bystanders using position trackers (VIVE Trackers) worn on hats, which provide position, and six-degree of freedom orientation information. While this method is relatively crude, we use this in place of indoor positioning systems which is a near-future technology (see [28, 31, 32] for reviews) beyond the scope of the present work.

In all cases, our visualizations operated in a two-state fashion: in the absence of bystanders, the visualization was not visible. When bystanders entered the tracking space (4m x 4m square—considered “social distance” [9]), the visualizations fade in to 50% transparency so the VR simulation always remains visible. When the Radar and Presence++ widgets are visible in this way, they are displayed in the periphery of the view (top left of the headsets-up display), as they were in their original incarnations [16]. The Radar and Presence++ widgets were placed 60 cm from the headset user, and at 60 degrees from the HMD’s forward vector. The center of the widget was placed 30cm to the left, and 20cm above the center of the headset user’s view.

5 USER AND SIMULATION HYBRID STUDY

We conducted a comparative study to understand whether the adapted designs sufficiently balanced the informal awareness of the physical environment with immersion in the virtual experience for the headset user. As described earlier, each visualization provides some informal awareness of bystanders in the environment (Table 1), but it is unclear whether the placement and abstractness of the information is appropriate, and whether it disrupts immersion. It may also be the case that how the information is presented should be subtle or more obvious. We were interested in several “bystander situations” that the headset user might find herself in: for example, whether there was one or multiple bystanders, and whether bystanders were interested in interaction, or simply passing through the space (Table 2). Participants experienced each visualization in three different VR conditions where bystander movements were pre-recorded and played back. The uniformity across participants is crucial in this study (e.g. how the bystanders moved, which ways they were positioned and oriented, etc.); had we used real actors (or other participants to behave as bystanders), it would have been much more challenging to ensure the bystanders behaved consistently for all participants. Simulating bystanders ensured that comparing participants’ reactions to the visualizations and conditions would be a fair and uniform.

The simulation study approach we adopt here is an increasingly common and accepted approach in the literature to study explore interactions—particularly to provide a stable form of

“bystander.” For instance, Mäkelä et al. study the impact of groups of simulated bystanders in VR (and find negligible differences in how people behaved around them) [27]. Similarly, Mahadevan et al. use simulated pedestrians in an exploration of how people behave around autonomous vehicles [25]. This is not an exhaustive list, but illustrate how prior literature have used bystander simulations in their work.

Our study therefore sought to address three related research questions: first, would participants feel they could maintain awareness of bystanders; second, would participants feel they could identify if and when bystanders wanted to interact with them, and finally, would the visualizations affect participants’ immersion into the VR scene? To study these questions, we designed a study where participants engaged in three VR experiences in a VR headset that showed the presence of bystanders. This study was approved by our local ethics committee.

5.1 Study Design

Participants. We recruited twenty participants (10 males and 10 females; mean age 22.4; $\sigma=1.80$) from our local university, numbered P1-P20. Twelve participants had prior experience with VR content, and the others did not have prior experience with VR content. None of them had any knowledge about this study before the experiment.

Visualization Conditions. We tested the three-visualization conditions described above (Avatar, Radar, Presence++). To understand how immersed people are in the VR content itself, we also used a baseline condition (where there is no visualization presented).

Our broad goal is to provide the headset user with continuous awareness of bystanders around the headset user, even if they do not want to interact with her [23]. In this study, we evaluate the techniques from two way. First, we explore how well each visualization can perform through quantitative and qualitative evaluation results. Second, we compared them to know what ways each method is better than the other through the results.

Bystanders. To simulate various bystander situations, we developed eight vignettes, each lasting roughly less than 30s, that represent common situations that occur within an office setting. We developed these vignettes based on our own experiences in a VR development lab—representative, but not necessarily covering every possible situation (e.g. high urgency). As described in Table 2 these vignettes varied in terms the number of bystanders that were part of the vignette, as well as the overall narrative of the vignette. Briefly, these varied from a single bystander passing through the space twice (e.g. to collect a book), to multiple bystanders meeting each other within the space and engaging in a conversation with one another, to bystanders that might stop with the goal of interacting with the headset user. In a given 5-minute trial, participants would experience all of the vignettes (plus a “control vignette”—i.e. a 25s vignette of no bystanders—the average time of the eight vignettes) in random order, where there was a random gap of 5s to 10s between vignettes. To enhance the believability of the bystander vignettes, we used furniture elements of the physical study environment (i.e. bystanders each had a desk, and the room had a bookshelf and two doors, as illustrated in Figure 4), and explained this relationship between the vignettes and the physical environment to the participants. Participants did not know how many vignettes would play.

VR Content Type. VR experiences vary in terms of their immersiveness—for example, how quickly a user needs to respond to the content, or to look around to engage with the content. Whereas prior work has typically only explored one type of VR experience or another (e.g. [10, 29, 43]), we sought to systematically study the performance of each visualization against three different VR experiences. *Highly interactive:* a first-person shooter game (Figure 5 (a)), where the participant is to shoot at balloons that randomly appear from around the player. In the first-person shooter game, we asked participants to shoot as many balloons as possible. During the game, participants could see their score (balloons popped) displayed on walls. *Moderately interactive:* a 3D sketching application (Figure 5 (b)), where the user was to sketch a free form sketching using a line-drawing tool. In the 3D sketching application, we gave participants themes

<i>Count</i>	<i>Interrupt</i>	<i>Description</i>
1	Y	Narrative: John walks into the room. He wants to invite the participant for lunch. He sees the participant is playing a VR game. He faces the participant from a distance, and waits.
1	Y	Narrative: Luis is working at his desk, and ultimately wants the participant to test his app. Luis works at his desk for a while. Then, he gets up and approaches to participant to ask for help.
1	N	Narrative: Ken walks into the room but forgets why he came in. He turns back to leave, but pauses at the door. He checks to make sure he has the right book in his hand, then leaves the room.
1	N	Narrative: Ken enters the room because he needs to grab something from John's desk. Ken goes straight to John's desk and looks in the desk for a while. Once he finds it, he leaves the room.
2	Y	Narrative: Kaori is working at her desk. Ken walks into the room. Kaori asks Ken for help. Ken goes to Kaori's desk for a chat. After a while, Ken turns to the participant to get her help, too. Ken approaches the participant to get her attention.
2	N	Narrative: Kaori walks into the room to work at her desk. She goes directly to her desk, and starts working. While Kaori is doing that, Leo is working at his desk.
4	Y	Narrative: Leo and Jack walk into the room since they are looking for a book. They walk around the room (near the headset user). They find a book on participant's desk. So, they approach to participant to ask the participant if it is okay to borrow it. While they are looking for the book, Ken enters the room, and talks with Kaori. Then, he leaves the room.
4	N	Narrative: Leo and Luis are working at their desks. Kaori and Jack walk into the room since they are looking for Kai. But Kai is not in the room. So, they leave the room.

Table 2. The bystander vignettes to involve 1-4 bystanders. In half, a bystander would attempt to interact with the participant.

before each trial, and they would draw something imagined from the theme. *Low interactive*: a 360° scenery movie, where the participant can freely look around. For the 360° scenery movie (Figure 5 (c)), we asked participants to watch the movie while thinking about where they would like to go. Participants watched a different movie in each visualization condition so they would not be bored. The movie presentation order was counterbalanced across visualization techniques for all participants. For every content type, participants stood, and played within a red circle drawn on the floor in the virtual environment. To focus the effect of the design variables, we kept the environment simple, even while varying the interactivity in the experience.

Method. Participants engaged in a total of 12 trials with the system (3 visualization conditions \times 3 content types plus 1 baseline condition \times 3 content types), where in each trial, the nine different “bystander vignettes” would play in random order. Prior to the experimental trials, participants engaged in a warm-up trials to ensure they understood how to engage with the VR content, and how each visualization worked. In each trial, the participant wore a headset and engaged in a dual task. In the primary task, the participant engaged with the VR contents (Figure 5). The participant's secondary task was to monitor bystanders who were visualized using one of the three visualization techniques, and then to signal (by pressing the trigger on the left controller) if she thought a bystander intended to speak with her. At the conclusion of each trial,

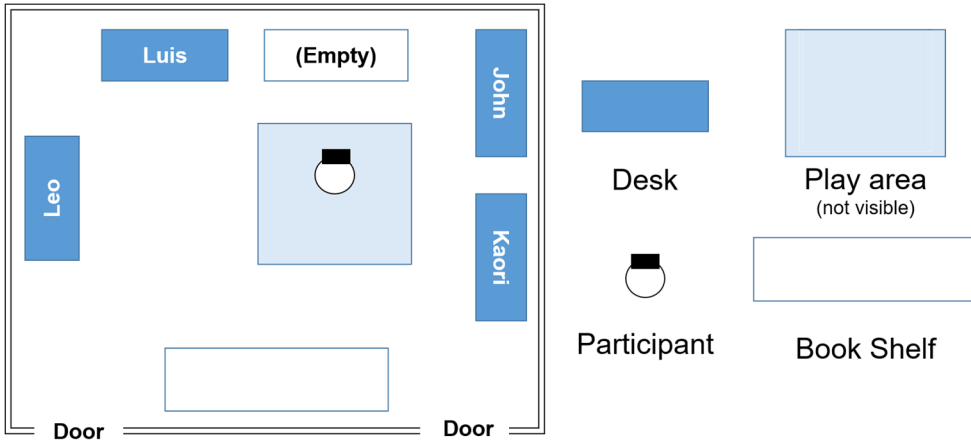


Fig. 4. Environment in vignettes.

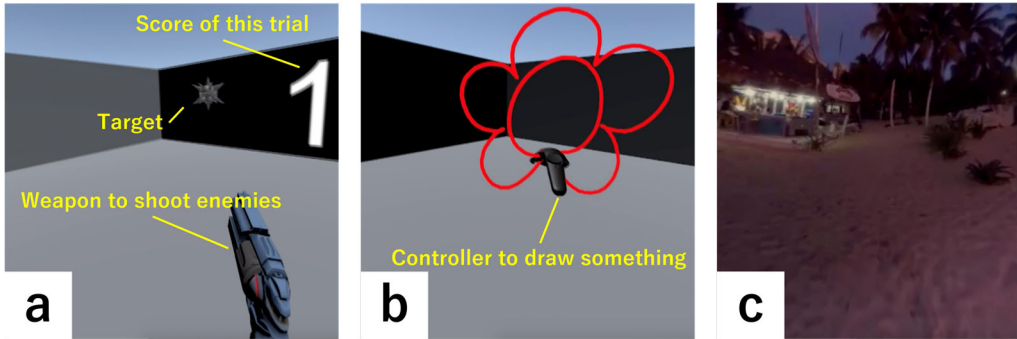


Fig. 5. We designed three different VR content types: (a) a first-person shooter game, (b) a 3D sketching, and (c) a 360-degree movie.

the participant responded to a questionnaire about the experience. At the conclusion, we conducted a semi-structured interview to understand participants' experiences. The ordering of the four visualization conditions (three visualizations + one baseline) was counterbalanced using a Latin square. The study took approximately 2.5 hours including breaks for each participant. In the baseline, no visualization condition, participants engaged in each VR content experience for five mins each without the secondary task—roughly the same duration as all of the nine bystander vignettes.

Measures. Our goal was to assess participants' immersion with the virtual simulation, their awareness of the bystanders, as well as their subjective report of the experience. To assess immersion, participants responded to a questionnaire based on the standard Igroup Presence Questionnaire (IPQ) [38] where we selected three of four sub-scale (general presence, spatial presence, and involvement), excluding the sub-scale focusing on experienced realism. We included two additional five Likert-scale questions to assess participants' awareness of bystanders in the environment: Was it easy to stay aware of the bystanders? Was it easy to identify the bystander who has an interest in you? Finally, we collected their performance on the bystander acknowledgement subtask (called identification score)—the number of times they correctly acknowledged a bystander that intended to interact with them (a maximum score of 4—the number of times a bystander interacts with the participants). In addition, we conducted a follow-

up interview to gather feedback from participants about the individual techniques, and their final preference ranking of each of the techniques.

System Design and Implementation. The test harness would run the VR experience, and in parallel, play back pre-recorded bystander movements (Table 2). Thus, the study simulated the presence and movements of bystanders (i.e. rather than using real bystanders for each participant). Participants were aware that the bystanders were simulated. Our prototypes were built using an HTC Vive headset and its controllers, and a wireless headphone (SONY, h.ear on 2 Wireless NC). The software ran on a desktop PC and controlled using Unity.

5.2 Results

Our study shows that the AVATAR view, as a situated, literal visualization of bystanders, was the best-performing as an awareness cue; however, PRESENCE++ preserved participants' immersion in the VR content the best across the visualization techniques. We briefly discuss the quantitative results before describing our qualitative findings. Since our experimental data was non-parametric, we used the Friedman test, a non-parametric analog to a repeated-measures ANOVA that detects differences between treatments, and used the Nemenyi test for post-hoc tests.

Were participants able to identify when bystanders wanted to interact with them? Based on our scenarios, participants needed to correctly identify four instances of when a bystander wanted to interrupt them per visualization condition and content type. Participants generally performed reasonably (overall=2.76, σ =1.20) on the IDENTIFICATION metric (max=4), and as illustrated in Figure 6 (left), performed best with the AVATAR visualization across all three content type conditions. Also, PRESENCE++ condition performed worst out of the three visualization conditions. The Friedman test found significant differences between visualization conditions ($p<.01$, $\chi^2=10.17$). A post-hoc Nemenyi test revealed significant difference between AVATAR (avg=3.18, σ =0.56) and PRESENCE++ (avg=2.10, σ =1.14) ($p<.05$, $r=0.67$). This effect was strongest in the FPS game and in the 360° movie. The Friedman test found significant differences between visualization conditions in FPS game ($p<.01$, $\chi^2=10.45$) and 360° movies ($p<.01$, $\chi^2=15.57$). A post-hoc Nemenyi test revealed significant differences in the FPS game between AVATAR (avg=3.65, σ =0.57) and PRESENCE++ (avg=2.35, σ =1.42) visualization conditions ($p<.05$, $r=0.65$), and in the 360° movie condition between RADAR (avg=3.40, σ =1.02) and PRESENCE++ (avg=1.80, σ =1.44) ($p<.01$, $r=0.64$), and between AVATAR (avg=3.40, σ =0.49) and PRESENCE++ ($p<.05$, $r=0.74$) (Figure 6 right). We analyzed effect of gender on this measure, and there was no impact.

Did participants feel like they could use the visualizations to maintain an awareness of the bystanders? Participants felt the visualizations helped them to maintain an awareness of bystanders. AVATAR and PRESENCE++ was got high score more than neutral, but RADAR was lower than it. This show that participants felt confident to maintain awareness when they were using AVATAR and PRESENCE++. Participants rated the AVATAR (avg=3.86; σ =0.79) visualization highest on this measure. The Friedman test found significant differences between visualization conditions ($p<.05$, $\chi^2=7.56$). A post-hoc Nemenyi test revealed significant difference between AVATAR and RADAR (avg=2.98; σ =0.81) ($p<.05$, $r=0.68$) (Figure 7). Here, participants did not vary significantly their rating of the visualization conditions depending on the content type; the exception to this is in the FPS game. The Friedman test found significant effect of the visualization conditions ($p<.01$, $\chi^2=17.48$) in FPS game. A post-hoc Nemenyi test revealed a significant difference between the AVATAR view (avg=4.30, σ =0.84) and the RADAR widget (avg=2.75, σ =1.26) ($p<.01$, $r=0.76$). We analyzed effect of gender on this measure, and there was no impact.

Did participants feel like they were able to identify bystanders who had an interest in them? Even though participants were objectively generally able to identify bystanders that wanted to interact with them, they did not feel consistently confident in using the techniques to do this. Only Avatar view got high score than neutral. Presence++ and Radar were lower score than neutral, and this show that participants did not feel confident when using these techniques.

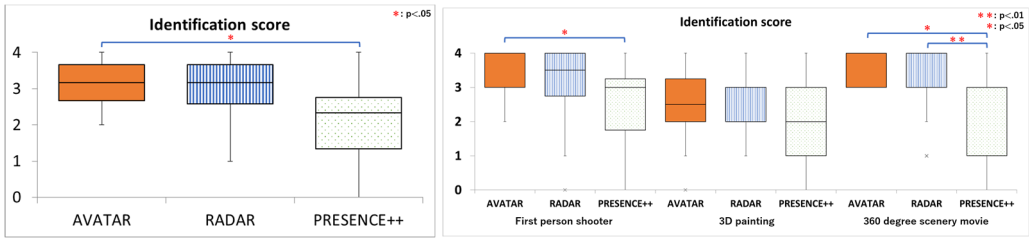


Fig. 6. Identification score overall (left), and identification score per content (right).

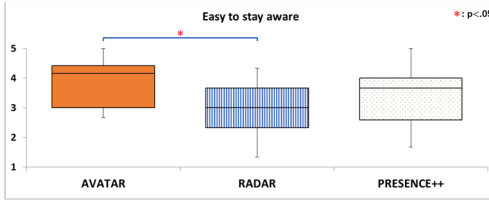


Fig. 7. Likert responses for: “Was it easy to stay aware of the bystanders?”

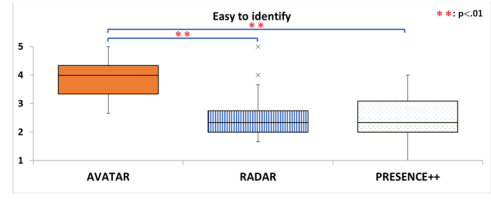


Fig. 8. Likert responses for: “Was it easy to identify the bystanders who had an interest to you?”

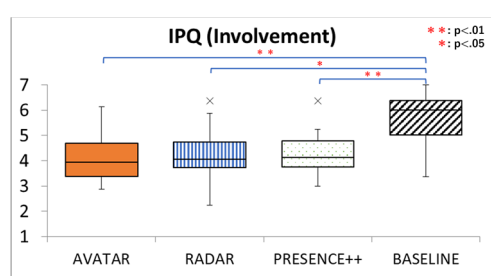
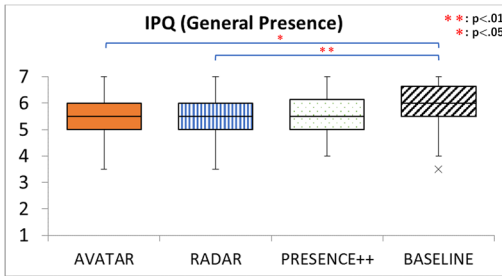


Fig. 9. IPQ results for each awareness technique.

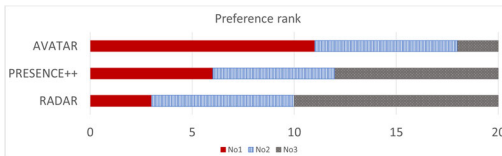


Fig. 10. Preference ranking of visualization condition

The Friedman test found significant effect difference between visualization conditions ($p < .01$, $\chi^2 = 20.24$). As illustrated in Figure 8, a post-hoc Nemenyi test revealed significant differences between the AVATAR condition (avg=3.90; $\sigma=0.72$) and the RADAR (avg=2.60; $\sigma=0.83$) ($p < .01$, $r=0.84$) and the PRESENCE++ conditions (avg=2.48; $\sigma=0.85$) ($p < .01$, $r=0.83$). We also analyzed effect of gender difference for the performance. A Wilcoxon signed-rank test revealed a significant impact of gender (Men: avg=3.10; $\sigma=0.87$, Women: avg=2.10; $\sigma=0.37$) in the RADAR condition ($p < .05$, $r=0.78$). In our study, men felt they could identify bystanders that wanted to interact with them using RADAR better than women.

How did the awareness visualizations affect immersion? Based on the ratings on the Igroup Presence Questionnaire (IPQ), PRESENCE++ rated highest in terms of being able to preserve

participants' immersion with the VR content across three visualization conditions. We explore three aspects of this: general presence, spatial presence and involvement. The three visualization conditions each scored higher than neutral for each aspect. In terms of the general presence, the PRESENCE++ (avg=5.60; σ =0.80) was higher than AVATAR (avg=5.48; σ =0.87) and RADAR (avg=5.45; σ =0.82) (Figure 9 left). The Friedman test found a significant difference between the visualization conditions ($p<.01$, $\chi^2=18.19$). A post-hoc Nemenyi test revealed that AVATAR ($p<.05$, $r=0.55$) and RADAR ($p<.01$, $r=0.58$) were significantly lower than baseline condition (avg=5.93; σ =0.95). In terms of the spatial presence, we did not find significant differences between the visualization conditions (PRESENCE++ (avg=4.99; σ =0.58), RADAR (avg=4.91; σ =0.40), AVATAR (avg=4.87; σ =0.50), baseline condition (avg=5.17; σ =0.54)). In terms of the involvement, the Friedman test found significant differences between visualization conditions ($p<.01$, $\chi^2=18.50$). A post-hoc Nemenyi test revealed that AVATAR (avg=4.12; σ =0.91) ($p<.01$, $r=0.74$), RADAR (avg=4.29; σ =0.98) ($p<.05$, $r=0.74$), and PRESENCE++ (avg=4.27; σ =0.98) ($p<.01$, $r=0.78$) were significantly lower than baseline condition (avg=5.63; σ =0.98) (Figure 9 right). Overall, this shows that the participants were more distracted by AVATAR and RADAR (where there is detailed information about the bystanders) compared than baseline condition. While we found that AVATAR and RADAR got a lower score than PRESENCE++, and that PRESENCE++ is better solution in terms of the immersion, all three visualization techniques produced lower immersion scores compared to the baseline condition. Finally, in our analysis of gender on these measures revealed no impact.

How did participants rank the visualizations? Participants ranked the AVATAR and PRESENCE++ visualization techniques the highest. As illustrated in Figure 10, the AVATAR received the most first and second place votes.

5.3 Findings from User Interviews

Our interviews results mirror the quantitative results: participants generally preferred the AVATAR view as it was easier for them to see and interpret; however, they were well aware of its weaknesses (inability to see behind themselves, and disruptive to the immersive experience). As a situated awareness cue, AVATAR was simply more straightforward to see, whereas the separate awareness widgets (RADAR and PRESENCE++) required participants to refocus their eyes away from the VR content when they wanted to see it. On the other hand, the AVATAR view presents large cues that were difficult to ignore and distracted from the main VR experience. In this respect, the PRESENCE++ cue was far less disruptive to the immersion.

What was the experience of the Separate awareness widgets work compared to the Situated awareness widget? The RADAR and PRESENCE++ widgets are displayed separately from the physical space in widgets that are placed 60 cm from the user. While this approach mimics the "heads up" view of many video games designed for 2D screens [44], and builds on significant prior literature on awareness widget design for collaborative work [16], the approach does not work as well in the VR context. Because of its placement (i.e. separate from the VR game environment, and in front of the game content), participants needed to move their eyes from the VR game to the awareness widget.

P12 describes having to do this shifting is problematic, and enough of a nuisance that they sometimes would forget to do it if the VR content is sufficiently engaging: *"To see the RADAR widget [or PRESENCE++], I need to move my eyes to the radar from VR content. When I focused on VR content, especially first-person shooter game and 3D painting, I forget to look at the RADAR, since I really focused on the content."* Similarly, P9 explains that they forget to do this when the content is immersive: *"When I focused on the VR content, especially in the FPS game and the Painting experience, I forgot to look at the RADAR altogether!"*

These experiences speak to two aspects of the RADAR and PRESENCE++ widgets: first, the user's eyes need to perform an accommodation when switching from the VR content (which tends to be "further away" from the user than the widget (which tends to be a bit closer); second, it reflects

the “foreground-only” design of the RADAR and PRESENCE++ widgets—that is, information cannot be obtained from seeing them in the periphery of user’s view, and so to get any information from the widget requires the user to switch their eye’s focus directly to the widget itself.

On the other hand, the Separate widgets did give participants awareness of what was happening behind themselves, which is something they appreciated and found missing in the AVATAR view. P2 describes this as a shortcoming of the AVATAR view: *“Sometimes, I was anxious since I couldn’t tell whether there was a bystander behind me or not.”* Similarly, P10 complained about the narrowness of the field of view of the AVATAR view, *“When I played 3D Painting game, sometimes I would keep facing forward. So, I may have missed bystanders; I don’t think I would have done this with the other visualization techniques.”*

P14 describes another reason why we take the separate widget. *“A bystander who will talk to you will probably come in front of [you]. So, the AVATAR view works well. But, the bystander who do not have interest to the headset user move around [without paying attention to what you are facing]. So, the headset user may miss [bystanders using] the AVATAR view.”* Also, P14 said that *“[Using radar or Presence++] I am able to know the bystander’s situation around me, and this help us to ensure safety”*.

Because situated awareness widgets are affixed to the locations of the bystanders as they would be in real life, they are easier to see, and it is easier to imagine what the bystander is doing. P11 reports, *“It’s easier to see bystanders [with AVATAR view], since they are within the center of the field of view.”* Thus, the participants felt that the AVATAR view was essentially like “seeing through” the HMD, and therefore interpreting bystanders’ actions and movements was much easier. P5 describes this feeling, *“[With AVATAR view], we can easily understand the bystanders’ behaviours very quickly. It’s easier to do this since it’s the same as in the physical world.”* P5 further explain this in contrast to the RADAR widget, which requires cognitively “remapping” what they see of the bystanders in the widget into the physical world, *“I need lots of time to understand the bystanders’ behaviours [with the RADAR widget] compared to other techniques, since I need to translate the RADAR information and imagine it in real space.”*

What was the experience of the Symbolic presentation style cue compared to the Literal presentation style? PRESENCE++ translates the bystanders’ behaviours into abstract symbols, basic information was easy to interpret quickly. For instance, P8 reports, *“It is easy to know presence of bystanders [with the PRESENCE++ widget],”* which is corroborated with P3’s comment, *“It is easier to see the information compared to the RADAR widget.”* It was also less disruptive: P16 reports, *“PRESENCE++ could provide awareness information around us, and I did not feel the cue disrupt the VR experience.”* Yet, because the abstraction deliberately obscures information about the bystanders (e.g. their location), or requires the user to remember an arbitrary mapping, it can be sometimes difficult to gain detailed information about the bystanders. For instance, P16 reported worrying that the abstraction might lead to incorrect conclusions: *“I may sometimes react to the bystanders who do not have interest to me.”* In addition, P17 reported that *“It was difficult to judge just by looking at them right away, so I had to decide whether or not to react by setting up criteria to judge who would talk to me and who would not through repeated experiences.”*

In contrast, the Literal presentation styles of the RADAR and AVATAR views sometimes provide too much information that overwhelms the user. Reporting on the AVATAR condition, P4 reports, *“I was confused [using the AVATAR view] when I was playing the FPS game, because I needed to distinguish between the bystanders and the enemies.”* Thus, although the AVATAR awareness widget was easier to interpret because it provided a more literal mapping of the bystanders’ positions, this interfered with the VR content. This was even more problematic for participants in the RADAR condition when there were multiple bystanders, *“The amount of information is a bit much [P5]”,* particularly since it shows so much information of the bystanders. Furthermore, movement in the awareness widget itself can be distracting: *“The arrows move around in the RADAR widget so frequently, it attracts my attention. [P7]”*

What aspects of the awareness widgets were useful and problematic? *AVATAR view:* Participants generally felt that the AVATAR view was effective in helping them understand when a bystander wanted to interact with them. The situated literal presentation means that to provide awareness it is disruptive to the visual field; however, the user can quickly make the decision to ignore it if necessary. P8 explains, “*AVATAR view is good for identifying bystanders that want to talk. So, AVATAR view can be good for urgent content, and a content which ask people to concentrate.*” The weak points with this widget was that the motion created by the bystanders was distracting, and user miss the AVATAR view if the AVATAR exist out of user’s field of view.

RADAR widget: The utility of the RADAR widget in its ability to provide information about bystanders around the user (even behind the user). But, because it was difficult to interpret and took time to really understand, participants felt it should be used only for content that generally did not require too much activity on the part of the user. P10 emphatically suggests, “*RADAR widget is suitable for low-interactivity content, such as a movie.*”

PRESENCE++ widget: The PRESENCE++ widget was effective in showing participants when bystanders were nearby, and less interfere with their immersion in the game. P10: “*Because I can easily confirm the presence of the icon, it is the same as knowing when bystanders are present.*” Its simplicity allows participants to easily ignore it, as reported by P12, “*I felt the PRESENCE++ widget preserves the VR experience better than the other techniques.*” The challenge is that it does not effectively present information that allows users to effectively know when someone wants to interact with them. Furthermore, because PRESENCE++ only showed the current state of bystanders, and it was difficult to recover the history of the bystanders (i.e. their past trajectory). Thus, it seems that Presence++ can be good for non-urgent situation (e.g. just someone entering a room) to provide awareness, while prioritize the immersion of VR experience.

Are continuous representations appropriate? In the absence of the visualizations altogether (i.e. in the baseline condition), participants were clear that the lack of awareness was problematic. For instance, P12 reported, “*During the tasks, I worry about my surroundings. Since I'd be embarrassed if I was moving my body in a weird way without knowing if anyone was there.*” Participants are aware of the fact that while they are in the immersion, they continue to physically exist in the real world. They moderate their actions when they know bystanders are present—even if the bystanders have no intention of interacting with them; therefore, it is important to have continuous representations since people mediate their interactions with others continuously—not discretely.

Limitations. While other studies have explored headset user-bystander interaction [29, 43], we used pre-recorded data to simulate bystanders, to ensure a consistent experience for each participant. This allowed us to explore subtle information (e.g. head orientation of bystanders) and situations with multiple interacting bystanders—all while ensuring that we retained high levels of experimental control (i.e. “bystanders” always behaved in the exact same ways). Though we can expect to get close results from the simulation study [27], this experimental control comes at a cost: we do not control the way in which bystanders may configure their behaviours to accommodate the headset user. For instance, a real-life bystander intending to interrupt the headset user might *deliberately* walk in a way that is directly in the headset user’s head orientation. This is not behaviour our study can capture; thus, participants’ scores in our study are likely an underestimate of their awareness of real-world bystanders.

In addition, participants knew there were no real bystanders in the physical space, which might have affected participants’ immersion, since the participants would not bump into bystanders even if participants were to collide with an avatar in virtual environment. During the task, the headset user only needed to press a button on the controller when they thought they were being “spoken to” by a bystander. But, we do not provide feedback about the “correctness” of this assessment; thus, participants cannot “learn” from their experience with the visualization techniques. On the other hand, the participants did not need to deal with the social consequences

of being “right” or “wrong” on their assessments. Thus, the scores we get in our study are “gut reaction” assessments of bystanders’ intentions.

Our study setup assumes headsets that entirely block sound. Others have explored more “current generation” technologies that provide visual immersion, but not full aural immersion (i.e. real-world sounds bleed through) [33, 34].

While we could design visualization cues more diegetic, or more realistic (e.g. [18, 43]), we designed simple cues that focused on the functional aspects. Because we would like to explore how the two design variables affect headset user’s awareness and immersion, we tried to keep other design variables simple. We acknowledged that if we design the cues from different variables, we may get different effects. For instance, [18] show five representation (indexical, symbolic, and iconic with three emotions), and the paper shows these cues had different effect in terms of awareness, and immersion. If we design our cues more indexical, the headset user can get rich information and this may improve awareness score.

Finally, we used simple VR environment for the tasks, and ask participants do the contents within particular zone. We acknowledge that if we use more complex VR content and allow participants to move around freely, these may affect the visualizations performance (e.g. spatial presence of the IPQ).

6 DISCUSSION AND FUTURE WORK

The core problem we address in this work is balancing informal awareness of bystanders with the disruptiveness to the VR experience for the headset user. Based on our study, we observed that each cue affected awareness of bystanders and immersion with the VR content differently. Our study revealed that surprisingly, the Radar widget performed consistently poorest. Based on our findings, we reflect on how to design effective awareness widgets in VR for headset users, and more broadly on how to design for awareness to permeate the HMD boundary.

Combining cues for balancing awareness and immersion. Our explorations show that each cue provides distinct benefits. The Avatar view, for instance, provides headset users with a straightforward way of understanding when a bystander is truly interested in interacting with them—bystanders are represented in a situated way so that their physical proximity with the headset user is easy to understand and interpret. Yet, it has weaknesses—it is very disruptive for the headset user’s experience (since the avatar is so large), and it cannot provide adequate awareness of bystanders that may be approach the headset user from behind. On the other hand, the Presence++ widget can provide an awareness of the bystanders’ presence without being as distracting because its presentation is separate from the VR scene itself. This type of peripheral, separate approach is powerful because it does not disrupt the immersive experience as much. Thus, designers should explore visualization methods and widgets that can convey information on the periphery of the headset user’s visual focus—that is, they should be able to glean information without having to visually attend to the information. Xiao et al., for example, use lighting cues along the edge of the head-mounted display that “glow” in various ways to provide peripheral awareness of out-of-view VR content [45]. We believe this approach, which is common 2D first person shooter games [6], can be re-deployed here to peripherally signal presence of bystanders.

Awareness need not be provided through a single widget; instead, it may be possible to present this information to users in multiple ways. For instance, depending on the level of urgency, different visualization techniques could be used—when bystanders are safely in the periphery, simply showing them through a Presence++ widget may be sufficient, and when they move closer into the headset user’s physical proximity, the Avatar view may be appropriate to ameliorate dangerous situations. Alternately, we might consider varying the opacity of the visualizations to preserve the VR experience for as long as possible. We intend to explore how to trigger and control these changes in the presentation of these kinds of awareness cues based on determinants

such as safety or urgency. For instance, if the headset user is about to collide with a bystander, then the visualization should be immediately visible and break the immersion. Similarly, some participants missed times when the bystander was talking to them—in these cases, we may need even stronger cues.

The converse may also be interesting to explore. In real life, bystanders can fade into the periphery of our attention if they do not move or we grow accustomed to their presence. In the same way, how can we slowly reduce the noticeability of visualizations if bystanders do not move for a long time (i.e. how do we allow them to fade into the background of the headset user's attention)?

Revealing additional awareness information about bystanders. In this research, we focused on tracked location and orientation information since they are important elements of informal awareness. However, as tracking technologies improve, we can soon collect additional information about the bystanders in the physical space (e.g. eye gaze direction, hand or arm movements and gestures)—e.g. through 360° cameras attached to the ceiling or headset [1]. This information can also be made available to headset user as part of that informal awareness (e.g. what is the bystander looking at? is the bystander gesticulating toward the headset user?), which can help the headset user make effective decisions about whether the bystander is trying to get her attention.

Awareness visualizations should be appropriate for and adapt to the virtual content. In our study, we showed that depending on the type of virtual content (e.g. fast-paced games vs. slower non-interactive content), different awareness visualizations might be appropriate; thus, awareness visualizations should be designed and chosen based on this type of consideration.

Some types of interactive content, such as first-person shooter video games, for instance, demand a lot of visual attention. To provide awareness, visualizations need to be equally demanding to get headset users' attention. In contrast, some content does not require constant visual attention (e.g. watching immersive VR scenery or a 360° video), so more subtle widgets may be more appropriate. Yet, VR content can be characterized along a number of other dimensions (e.g. "high interactivity" can come in many forms—whether it involves considerable moving around, or turning one's head, or moving one's arms)—and these other dimensions may also impact the particular design choice for such visualizations.

Similarly, visualizations should be adaptable to the virtual environment itself. In our study, the virtual environments were simple, but if the virtual environment was visually dense and crowded (e.g. with lots of objects or walls), it may be the case that the awareness visualization might not be as visually salient (or disrupt the experience too much). Minimally, designers need to consider the colors of the visualization in relation to the environment; more broadly, designers will likely need to explore multi-modal techniques (e.g. sound or haptics), content-based design, or emotion [12, 26]. For instance, O'Hagan et al. [33] show that when headset users employ open-ear headsets, they can maintain this awareness of bystanders to some extent already.

Providing informal awareness of the headset user's activities to bystanders. Our work specifically only deals with the concerns of headset users; however, because initiating interaction is a bi-directional, negotiated set of actions [5], designers should also provide bystanders awareness information about the headset user's activities. These motivations were explored by prior work [34], and would users in smoothly moving toward interaction: if bystanders understand what headset users are experiencing, they can modulate how they approach the headset user. Previously, others have proposed techniques to enable bystanders to explicitly interact with headset users [3, 14, 15]. In our work, we did not specifically study interaction by direct contact (e.g. a bystander tapping on the headset user's shoulder), as we would expect this would be extremely obtrusive. Our future interest is therefore in exploring how to present similar kinds of information to bystanders to enable them to smoothly move into interaction with the headset users as well.

7 CONCLUSION

Current VR experiences emphasize immersion without considering how to support informal awareness of bystanders sharing the physical space. Our work shows that it is possible to reach reasonable compromises in awareness visualization designs for VR that can provide some awareness of bystanders without overly disrupting the immersive experience. Based on our comparative study, we show that the Avatar view was the most effective in providing awareness of bystanders, but that Presence++, which uses a more peripheral visualization style, was able to better preserve immersion. Thus, we recommend that VR designers consider the use of peripheral visualizations that do not demand foreground visual attention for displaying informal awareness cues, and to explore how to visualize multiple levels of urgency. These awareness signals help mitigate the HMD boundary, which will allow headset users to maintain an awareness of their surroundings. This awareness will reduce awkward interactions across the HMD boundary, returning agency to headset users over their engagement with nearby bystanders.

ACKNOWLEDGMENTS

This work was supported in part by JSPS Grants KAKENHI (19KK0258).

REFERENCES

- [1] Jérôme Ardouin, Anatole Lécuyer, Maud Marchal, Clément Riant, and Eric Marchand. 2012. FlyVIZ: A Novel Display Device to Provide Humans with 360 ° Vision by Coupling Catadioptric Camera with HMD Categories and Subject Descriptors. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology (VRST '12)*: 41–44. <https://doi.org/10.1145/2407336.2407344>
- [2] Pulkit Budhiraja, Rajinder Sodhi, Brett Jones, Kevin Karsch, Brian Bailey, and David Forsyth. 2015. Where's My Drink? Enabling Peripheral Real World Interactions While Using HMDs. *arXiv:1502.04744*
- [3] Liwei Chan and Kouta Minamizawa. 2017. FrontFace: Facilitating Communication Between HMD Users and Outsiders Using Front-Facing-Screen HMDs. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*: 1–5. <https://doi.org/10.1145/3098279.3098548>
- [4] T. Matthew Ciolek and Adam Kendon. 1980. Environment and the Spatial Arrangement of Conversational Encounters. *Sociological Inquiry* 50, 3–4: 237–271. <https://doi.org/10.1111/j.1475-682X.1980.tb00022.x>
- [5] Herbert H. Clark. Using Language. 1996. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620539>
- [6] Kody R. Dillman, Terrance Tin Hoi Mok, Anthony Tang, Lora Oehlberg, and Alex Mitchell. 2018. A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*: 1–12. <https://doi.org/10.1145/3173574.3173714>
- [7] Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM Conference on Computer Supported Cooperative Work (CSCW '92)*: 107–114. <https://doi.org/10.1145/143457.143468>
- [8] Paul Dourish and Sara Bly. 1992. Portholes: supporting awareness in a distributed work group. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*: 541–547. <https://doi.org/10.1145/142750.142982>
- [9] E.T.Hall. 1966. Hidden dimension. Garden city, N.Y., Doubleday.
- [10] Robert S. Fish, Robert E. Kraut, and Barbara L. Chalfonte. 1990. The VideoWindow system in informal communication. In *Proceedings of the 1990 ACM Conference on Computer Supported Cooperative Work (CSCW '90)*: 1–11. <https://doi.org/10.1145/99332.99335>
- [11] Ceenu George, Philipp Janssen, David Heuss, and Florian Alt. 2019. Should I interrupt or not? Understanding interruptions in head-mounted display settings. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*: 497–510. <https://doi.org/10.1145/3322276.3322363>
- [12] Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander, Jeff Hotnig, Douglas Cheong, Gabriel Reyes, and Gregory. D. Abowd. 2018. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4: 1447–1456. <https://doi.org/10.1109/TVCG.2018.2793698>
- [13] Saul Greenberg. 1996. Peepholes: Low Cost Awareness of One's Community. In *Conference Companion on Human Factors in Computing systems (CHI '96)*: 206–207. <https://doi.org/10.1145/257089.257283>
- [14] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*: 4021–4033. <https://doi.org/10.1145/3025453.3025683>

- [15] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*: 1–13. <https://doi.org/10.1145/3173574.3173628>
- [16] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Comput. Supported Coop. Work* 11, 3: 411–446. <https://doi.org/10.1023/A:1021271517844>
- [17] Jeremy Hartmann, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. RealityCheck: Blending virtual environments with situated physical reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*: 1–12. <https://doi.org/10.1145/3290605.3300577>
- [18] Zhenyi He, Fengyuan Zhu, Ken Perlin, and Xiaojuan Ma. 2018. Manifest the invisible: Design for Situational Awareness of Physical Environments in Virtual Reality. *arXiv:1809.05837*
- [19] Akira Ishii, Masaya Tsuruta, Ippei Suzuki, Shuta Nakamae, Junichi Suzuki, and Yoichi Ochiai. 2019. Let Your World Open: CAVE-based Visualization Methods of Public Virtual Reality towards a Shareable VR Experience. In *Proceedings of the 10th Augmented Human International Conference 2019 (AH '19)*: 1–8. <https://doi.org/10.1145/3311823.3311860>
- [20] Ikuo Kamei, Changyo Han, Takefumi Hiraki, Shogo Fukushima, and Takeshi Naemura. 2020. CoVR: Co-located Virtual Reality Experience Sharing for Facilitating Joint Attention via Projected View of HMD Users. In *SIGGRAPH Asia 2020 Emerging Technologies (SA '20)*: 1–2. <https://doi.org/10.1145/3415255.3422883>
- [21] HyeonYeop Kang and JungHyun Han. 2020. SafeXR: alerting walking persons to obstacles in mobile XR environments. *Vis Comput* 36: 2065–2077. <https://doi.org/10.1007/s00371-020-01907-4>
- [22] Adam Kendon. 2010. Spacing and orientation in co-present interaction. In *Proceedings of the Second international conference on Development of Multimodal Interfaces: active Listening and Synchrony (COST'09)*. Springer-Verlag, Berlin, Heidelberg: 1–15. https://doi.org/10.1007/978-3-642-12397-9_1
- [23] Jarrod Knibbe, Jonas Schjerlund, Mathias Petráus, and Kasper Hornbæk. 2018. The Dream is Collapsing: The Experience of Exiting VR. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*: 1–13. <https://doi.org/10.1145/3173574.3174057>
- [24] Robert E. Kraut, Robert S. Fish, Robert W. Root, and Barbara L. Chalfonte. 1990. Informal Communication in Organizations - Form, Function, and Technology. I. S. Oskamp & S. Spacapan (Eds.). *Human Reactions to Technology: The Claremont Symposium on Applied Social Psychology*, Beverly Hills, CA: Sage Publications.
- [25] Karthik Mahadevan, Elaheh Sanoubari, Sowmya Somanath, James E. Young, and Ehud Sharlin. 2019. AV-pedestrian interaction design using a pedestrian mixed traffic simulator. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*: 475–486. <https://doi.org/10.1145/3322276.3322328>
- [26] Christian Mai, Mariam Hassib, and Ceenu George. 2017. Like Elephants Do: Sensing Bystanders During HMD Usage. In *Proceedings of the CHI 2017 Workshop on Amplification and Augmentation of Human Perception*.
- [27] Ville Mäkelä, Rivu Radiah, Saleh Alsherif, Mohamed Khamis, Chong Xiao, Lisa Borchert, Albrecht Schmidt, and Florian Alt. 2020. Virtual Field Studies: Conducting Studies on Public Displays in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*: 1–15. <https://doi.org/10.1145/3313831.3376796>
- [28] Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. 2011. The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)*: 315–326. <https://doi.org/10.1145/2047196.2047238>
- [29] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen. Brewster. 2015. 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 (CHI '15)*: 2143–2152. <https://doi.org/10.1145/2702123.2702382>
- [30] Daniel Medeiros, Rafael dos Anjos, Nadia Pantidi, Kun Huang, Maurício Sousa, Craig Anslow, and Joaquim Jorge. 2021. Promoting Reality Awareness in Virtual Reality through Proxemics. *2021 IEEE VR Virtual Reality and 3D User Interfaces (VR)*: 21–30. <https://doi.org/10.1109/VR50410.2021.00022>
- [31] Germán Martín Mendoza-Silva, Joaquín Torres-Sospedra, and Joaquín Huerta. 2019. A Meta-Review of Indoor Positioning Systems. *Sensors* 19, 20: 1–45. <https://doi.org/10.3390/s19204507>
- [32] Raul Montoliu, Emilio Sansano, Arturo Gascó, Oscar Belmonte, and Antonio Caballer. 2020. Indoor Positioning for Monitoring Older Adults at Home: Wi-fi and BLE Technologies in Real Scenarios. *Electronics* 9, 5: 1–20. <https://doi.org/10.3390/electronics9050728>
- [33] Joseph O'Hagan and Julie R. Williamson. 2020. Reality aware VR headsets. In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays (PerDis '20)*: 9–17. <https://doi.org/10.1145/3393712.3395334>
- [34] Joseph O'Hagan, Julie R. Williamson, and Mohamed Khamis. 2020. Bystander interruption of VR users. In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays (PerDis '20)*: 19–27. <https://doi.org/10.1145/3393712.3395339>
- [35] Daniel Pohl and Carlos Fernandez De Tejada Quemada. 2016. See what I see: Concepts to improve the social acceptance of HMDs. *2016 IEEE Virtual Reality (VR)*: 267–268. <https://doi.org/10.1109/VR.2016.7504756>
- [36] Anastasia Ruvimova, Junhyeok Kim, Thomas Fritz, Mark Hancock, and David C. Shepherd. 2020. “Transport Me Away”: Fostering Flow in Open Offices through Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*: 1–14. <https://doi.org/10.1145/3313831.3376724>
- [37] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task, and Environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19)*: 199–211. <https://doi.org/10.1145/3311350.3347190>

- [38] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and virtual environ.* 10, 3: 266–281. <https://doi.org/10.1162/105474601300343603>.
- [39] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*: 1–12. <https://doi.org/10.1145/3290605.3300590>
- [40] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*: 182–195. <https://doi.org/10.1145/3379337.3415827>
- [41] Chiu-Hsuan Wang, Seraphina Yong, Hsin-Yu Chen, Yuan-Syun Ye, and Liwei Chan. 2020. HMD Light: Sharing In-VR Experience via Head-Mounted Projector for Asymmetric Interaction. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*: 472–486. <https://doi.org/10.1145/3379337.3415847>
- [42] Steve. Whittaker, David. Frohlich, and Owen. Daly-Jones. 1994. Informal workplace communication: What is it like and how might we support it? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*: 131–137. <https://doi.org/10.1145/191666.191726>
- [43] Julius Von Willich, Markus Funk, Florian Müller, Karola Marky, Jan Riemann, and Max Mühlhäuser. 2019. 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)*: 487–496. <https://doi.org/10.1145/3322276.3322334>
- [44] Jason Wuertz, Sultan A. Alharthi, William A. Hamilton, Scott Bateman, Carl Gutwin, Anthony Tang, Zachary Toups, and Jessica Hammer. 2018. A Design Framework for Awareness Cues in Distributed Multiplayer Games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*: 1–14. <https://doi.org/10.1145/3173574.3173817>
- [45] Robert Xiao and Hrvoje Benko. 2016. Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*: 1221–1232. <https://doi.org/10.1145/2858036.2858212>
- [46] Soojeong Yoo, Phillip Gough, and Judy Kay. 2020. Embedding a VR Game Studio in a Sedentary Workplace: Use, Experience and Exercise Benefits. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*: 1–14. <https://doi.org/10.1145/3313831.3376371>
- [47] André Zenner, Marco Speicher, Sören Klingner, Donald Degraen, Florian Daiber, and Antonio Krüger. 2018. Immersive Notification Framework: Adaptive & Plausible Notifications in Virtual Reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*: 1–6. <https://doi.org/10.1145/3170427.3188505>
- [48] WhatsApp. <https://www.whatsapp.com/>
- [49] Messenger. <https://www.messenger.com/>

Received February 2021, revised May 2021, accepted June 2021.