"It Feels Like I am Invited to Communicate": Mediating Ad-Hoc Bystander-VR User Interruptions Through Proactive Proxies

Adil Rahman
Department of Computer Science
University of Virginia
Charlottesville, Virginia, USA
adil@virginia.edu

Wen Ying
Department of Computer Science
University of Virginia
Charlottesville, Virginia, USA
wy7yv@virginia.edu

Md Aashikur Rahman Azim
Department of Computer Science
University of Virginia
Charlottesville, Virginia, USA
ma6zp@virginia.edu

Michelle Annett
MishMashMakers
Newmarket, Ontario, Canada
michelle@mishmashmakers.com

Seongkook Heo*
Department of Computer Science
University of Virginia
Charlottesville, Virginia, USA
seongkook@virginia.edu

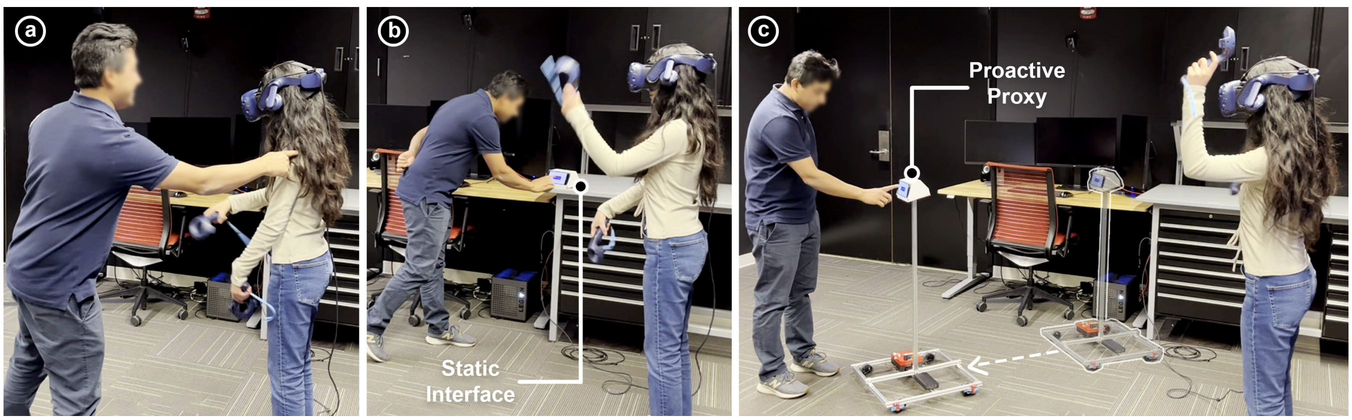

Figure 1: When someone is using VR, it can be difficult for a bystander to get their attention because (a) physical touch between strangers may be uncomfortable or (b) static interruption interfaces such as doorbells assume that bystanders are aware of them. (c) Proactive proxies, such as mobile robots with an interruption interface on them, however, may facilitate easier interruptions, even for uninformed bystanders.

Abstract

As VR expands into public spaces, new challenges emerge around spontaneous interactions between bystanders and unfamiliar VR users. While current VR systems often prioritize user awareness of their physical surroundings, they overlook the social dynamics affecting nearby bystanders. We conducted a deception-based study (N=80) examining how interface availability influences bystanders' comfort, confidence, and hesitation when interrupting VR users. We compared traditional static interruption interfaces (e.g., button on screen) with a proactive proxy that actively approached bystanders upon detecting interruption intent. Static interfaces, due to insufficient cueing, frequently caused bystander discomfort, leading to hesitant physical interruptions or complete communication avoidance. In contrast, the proactive proxy implicitly conveyed social permission, significantly enhancing bystanders' comfort and confidence. Our findings provide empirical insights into how bystanders assess availability and initiate interruptions with unfamiliar VR users in shared spaces, offering design implications for VR systems that support bystander agency and comfort during these interactions.

CCS Concepts

• **Human-centered computing → Virtual reality**; *Empirical studies in collaborative and social computing.*

Keywords

Virtual Reality; Bystander-VR User Interaction; Interruptions; Interruption Interface; Proactive Proxy

ACM Reference Format:

Adil Rahman, Wen Ying, Md Aashikur Rahman Azim, Michelle Annett, and Seongkook Heo. 2026. "It Feels Like I am Invited to Communicate":

*Corresponding Author

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3791936

Mediating Ad-Hoc Bystander-VR User Interruptions Through Proactive Proxies. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3772318.3791936

1 Introduction

Human interruptions in shared spaces are inevitable - whether it is asking a colleague for their input on a shared project [80] or asking a fellow passenger to move [81]. In these moments, gaining someone's attention is essential. However, interruptions become more challenging when individuals are deeply immersed in activities that render their primary communication channels unavailable (e.g., listening to music occupies their hearing, reading occupies their vision, etc.). In such cases, secondary communication channels like peripheral vision and ambient awareness serve as subtle cues for initiating interaction.

When individuals are immersed in virtual reality (VR), this dynamic becomes especially complex because their audiovisual immersion effectively limits the secondary use of these primary communication channels [19, 28, 40, 53]. Several alternative modalities have thus been proposed to manage interruptions in VR [50]. Modern VR headsets, such as the Meta Quest and Apple Vision Pro, incorporate open-ear audio technology that enables users to hear verbal interactions. However, this technology can reduce immersion and may not align with all users' preferences [50]. Additionally, such technology can be problematic in shared spaces, as sound leakage can disturb nearby individuals and make VR users self-conscious about audio spillover [50, 63]. Users already seek full auditory isolation when completing focused work in public environments by using noise-canceling headphones [46, 55], and as portable VR headsets become productivity tools, extending this need for isolation to the visual domain is a natural progression. Both the Apple Vision Pro and Meta Quest now include dedicated travel modes for use on airplanes and trains [4, 41], explicitly supporting fully immersive use in shared spaces. Moreover, airlines have also begun providing VR headsets to passengers as part of in-flight entertainment services [16], while academic and public libraries increasingly offer VR headset lending programs and dedicated VR spaces [2, 34]. These developments signal that encounters between unfamiliar individuals (i.e., one immersed in VR and the other needing to interrupt them) will become increasingly common. Critically, because bystander awareness features within headsets remain user-controlled, bystanders have no guaranteed means of initiating contact when VR users opt for full immersion. Touch, while potentially the most direct way to interrupt a VR user, is often socially inappropriate in public settings, where interactions among strangers are governed by implicit boundaries [50, 53]. As VR headsets transition beyond private environments into shared spaces where bystanders may need to interrupt VR users, developing socially appropriate and effective interruption mechanisms becomes increasingly critical [1, 14, 62, 72, 81].

One approach to this challenge has been to enhance VR users' awareness of their surroundings. Previous work has explored embedding awareness systems within VR headsets to help users avoid physical accidents and maintain spatial boundaries [22, 40, 42, 49, 51, 53, 58, 73, 84]. Such systems, however, cannot adequately replicate the dynamics of face-to-face interruptions and the subtle social cues that initiators (i.e., bystanders) carefully weigh when determining whether an interruption is appropriate [44]. By shifting control from initiators to VR users, awareness systems invert the natural negotiation process that governs interpersonal interruptions. Beyond immersion trade-offs and the risk of information overload in public spaces, embedded awareness systems may also prematurely redirect a VR user's attention to bystanders, triggering unwanted *face engagements* [20]. Although awareness systems can dynamically adjust how they deliver information by withholding alerts until a verbal interaction is attempted [52], such techniques overlook how bystanders assume others' unavailability and make decisions not to engage [76, 80]. In crowded public spaces, VR users may even disable these systems to avoid constant notifications about nearby bystanders, especially when they do not anticipate needing to interact with others.

Although limited, prior work has explored bystander-initiated interruption interfaces, such as the HTC's Knock Knock feature [71] and a physical doorbell peripheral [81], which offer explicit mechanisms for bystanders to attract a VR users' attention. However, these interfaces typically assume that bystanders are aware of their existence, limiting their effectiveness in public spaces where interruptions are often spontaneous and involve unacquainted individuals. While interruptions between acquaintances may be more frequent, stranger-to-stranger interactions represent the most challenging case due to the absence of established social rapport and the heightened psychological cost of initiating contact [12]. Solutions effective in such demanding contexts are also likely to generalize to less demanding scenarios. While prior research underscored the importance of observing spontaneous interactions between unfamiliar individuals [50], little is known about how such encounters unfold when unacquainted VR users share physical spaces. Yet, insights into these dynamics are essential for designing VR headsets that integrate more seamlessly into public environments. In this work, we investigate how bystanders naturally navigate spontaneous interactions with unfamiliar VR users in shared environments.

As studying such unplanned encounters at scale is challenging due to their situational constraints, we conducted a deception-based study that recreated these constraints in a controlled setting. Unlike prior work that relied on solicited interruptions [17, 23, 50], acquainted participants [50, 53], or anecdotal reports [53], our study placed 80 unsuspecting bystanders in scenarios requiring urgent interaction with an unfamiliar VR user in a shared space. Our study had two interruption interface conditions. The baseline condition, experienced by 40 participants, used a typical desktop VR setup with a static interruption interface (e.g., doorbell peripheral). For the second condition, we hypothesized that explicitly visible, bystander-initiated interfaces could improve interruption experiences. Inspired by public display research [27, 30], we created a robot-based interruption interface, i.e., a proactive proxy that served as a **design probe** to examine whether explicit interface discoverability impacted bystander experiences. This deliberately extreme intervention was intended to isolate the effect of proactive presentation on bystander comfort and behavior. This condition was experienced by an additional 40 participants.

During our study, we found that, similar to public kiosk research [30], the participants who encountered the baseline condition suffered from the *first-click problem* [30], where they failed to notice or use the static interface. In contrast, the participants who encountered the proactive proxy reported increased comfort and engagement. While participants preferred mediated interaction over direct interruption, the static interface was largely ignored. These findings underscore the importance of interface visibility and design in shared VR spaces, offering actionable insights for future VR headset development that better supports both users and bystanders.

This research contributes:

(1) Empirical evidence that bystanders feel significant discomfort during spontaneous interactions with unfamiliar VR users in shared spaces, with key insights revealing how physical boundaries, perceived safety, and social dynamics influence their interruption behaviors.
(2) Identification of key limitations in existing static interruption interfaces, showing how privacy concerns and low situational awareness hinder their discoverability and usability during spontaneous interactions.
(3) Design implications for bystander-aware VR interfaces that support comfortable and effective spontaneous interactions, emphasizing the need to explicitly integrate bystander perspectives into VR system design.

2 Related Work

Human interruptions are complex social behaviors that rely on various cues and coordinated actions between participants [10, 39, 57]. People engage in communication acts that go beyond verbal exchanges, incorporating subtle nonverbal elements such as body movements, physical proximity, and spatial orientation [57, 61]. For example, an airplane passenger leaning forward may suggest they wish to leave their seat [81]. In VR, however, these social cues are lost, as users' visual and auditory attention is absorbed by head-mounted displays. Their absence can worsen the psychological cost of interruption, as bystanders must weigh their need to interrupt against the risk of disrupting another person's ongoing task [12, 44, 61]. This may hint at why bystanders would hesitate to disrupt a user's immersion, choosing instead to delay or avoid interruption, as observed in some VR studies [17, 50].

2.1 System Approaches to Supporting Interruptions in VR

Prior research on the social acceptability of VR underscored the importance of supporting communication between the virtual and physical worlds [14, 81]. Much of this work has focused on awareness systems designed to alert VR users to nearby bystanders, using visualizations [18, 22, 37, 40, 42, 49, 53, 58, 73, 84] or incorporating external sounds into the VR environment [49, 51]. For example, Apple Vision Pro includes a feature called *EyeSight*, which enables partial pass-through to visualize bystanders during interruptions [3]. While effective at increasing awareness, these systems introduce a trade-off between maintaining immersion and providing sufficient awareness of the physical environment [52], as visual representations of bystanders can disrupt immersion [40] and increase cognitive load [56], while subtle notifications risk being overlooked [58]. Moreover, by shifting control from bystanders to VR users, awareness systems invert the natural negotiation of interruptions, sometimes preemptively triggering interactions that undermine bystander agency [17].

Beyond user-focused awareness, several systems have been developed to give bystanders explicit tools to attract attention, such as the HTC Vive's Knock Knock [71] or doorbell-like peripherals [68, 81, 84]. These approaches, however, often assume that bystanders possess prior knowledge of how to operate such interfaces or interpret the signals presented by VR users. O'Hagan et al. identified a similar challenge when survey participants were asked to evaluate different interruption interfaces [50].

Taken together, existing systems either compromise immersion (awareness-focused) or rely on bystander knowledge (bystander-initiated), leaving a gap in designs that balance immersion with bystander agency. This motivates further research into approaches that explicitly consider the needs of both VR users and nearby bystanders.

2.2 Empirical Understandings of VR Interruptions

While technical systems provide mechanisms for supporting interruptions in VR, less is known about how interruptions actually unfold in real social contexts. Some studies have investigated communication approaches between VR users and bystanders, such as sending notifications to VR users [58] or visualizing a bystander's avatar [35]. However, these have largely examined the VR user's perspective, while bystanders' experiences remain less understood. Many studies simulate interruptions by instructing confederates to act as bystanders or by prompting participants to respond to scripted events [17, 50]. These designs limit realism by constraining the spontaneity and uncertainty that characterize actual interruptions.

Another limitation is that most prior work has focused on acquainted pairs or solicited interruptions, where the bystander is known to the VR user or expected by design [17, 50]. In some studies, interruptions were explicitly elicited by participants who were aware of their role in the study [17, 58]. This awareness likely diminished the urgency and realism of the interruption scenarios, potentially obscuring the dynamic and context-sensitive factors that influence bystander behavior. Moreover, when bystanders were acquaintances of the VR user or expected by design, the psychological cost of interruption was reduced, overlooking how unfamiliar bystanders negotiate entry into an ongoing immersive experience. With the rapid development of VR technology, users increasingly wear headsets in public environments, where bystanders may be unfamiliar with them [81]. It is therefore crucial to understand how bystanders attempt to communicate with VR users they do not know. Studies that do consider the public use of VR and interactions between strangers often rely on anecdotal evidence gathered from VR users via surveys [50, 53] or through simulated scenarios [14, 62, 72]. While these provide useful perceptions, they cannot capture how people act in urgent or high-stakes situations.

As a result, our understanding of how interruptions involving strangers unfold and how they impact bystanders in public VR settings remains limited. This gap motivated the design of our study, which uses a deception-based methodology to observe authentic, stranger-to-stranger interruptions in situ, without the constraints of prior assumptions about the bystanders' knowledge or their relationship with the VR user.

2.3 Robot-Mediated Interactions

Inspired by research on public displays that use explicit presence to attract attention and support interaction [27, 30], we consider how robots can serve as proactive proxies for bystanders in VR interruption scenarios.

Due to their decreasing cost and increasing accessibility [82], robots have been deployed in a variety of public service roles across spaces such as shopping malls [7, 21, 31], academic conferences [25, 45], airports [69], supermarkets [8, 77], and company workspaces [33]. While interacting with robots in these settings has become increasingly socially acceptable [9, 70, 78], concerns remain about the interruptibility of such robots, particularly how and when humans can appropriately engage with them [47, 54, 59].

Prior research has examined interruptions from two primary perspectives: robots interrupting bystanders [9, 54], and bystanders interrupting robots [47, 59]. However, this leaves a gap in understanding a third possibility, i.e., a bystander using a robot as a proxy to interrupt another person in a public space. This third scenario is the focus of the present research, as service robots possess inherent affordances that make them highly noticeable to bystanders. For example, because human vision is acutely sensitive to even subtle environmental movements [25], any motion generated by a service robot tends to stand out in its surroundings. Moreover, motion has been shown to enable robots to convey emotions and intentions through non-verbal cues [25, 38, 61].

Despite the widespread use of robots in public spaces, their application in VR has primarily focused on enhancing user immersion through haptic environments [65, 66, 74, 75]. Some research has also explored the use of embodied robots for social interaction within VR [29]. In our work, we bridge these two lines of research (i.e., service robots in public spaces and robots in VR environments), by introducing the *proactive proxy*, a service robot designed to facilitate external communication with VR users. By leveraging the robot's inherent ability to attract attention, we extend its role to support more effective and intuitive interactions between VR users and bystanders.

3 Deception Study

To elicit natural reactions during spontaneous interactions with a VR user, participants were kept unaware of the study's true purpose until the session concluded. We employed a deception-based methodology, initially presenting the study as an investigation into cognitive and physical load during math problem-solving. Meanwhile, we covertly engineered a social situation in which the bystander (i.e., the participant) was prompted to engage with a VR user (i.e., an experimenter) via either a traditional static interruption interface or a proactive interface that explicitly presented itself. The study design ensured that (i) the bystander and VR user were complete strangers, (ii) the bystander experienced a sense of urgency to initiate interaction, and (iii) the bystander had no prior knowledge of the available interruption methods.

Based on our study design and findings from prior work, we hypothesized that:

H1: Bystanders will experience significant discomfort when initiating spontaneous interactions with unfamiliar VR users in shared spaces, which will manifest as hesitation or the complete avoidance of interruption behaviors.

H2: Traditional static interruption interfaces, even when clearly visible and logically positioned, will not effectively attract bystanders' attention during spontaneous interaction, resulting in significantly lower usage compared to proactive interfaces.

H3: Proactively presenting interruption interfaces will significantly reduce bystander discomfort and hesitation, thereby increasing both the likelihood and perceived quality of spontaneous interactions with unfamiliar VR users in shared spaces.

3.1 Participants

Eighty participants (40 female, 40 male; *Mean* = 22 years, *SD* = 3 years; *Range* = 18 - 30 years) were recruited to serve as bystanders during our study. They were recruited via university mailing lists and word of mouth. As a result, the study population primarily consisted of culturally and geographically diverse undergraduate and graduate students enrolled at a North American university. To balance the gender distribution across study configurations, an equal number of male and female participants[1] were recruited. The recruitment text did not mention VR interaction or the interruption of strangers. Instead, it informed potential participants that the study aimed to *"examine the physical and cognitive load of users while they performed various tasks"*.

Fifty-one participants had never used VR, 18 had used it occasionally, 9 used it sometimes, and 2 used it often. Fifty-nine participants had never interacted with people using VR, 16 had done so occasionally, and 5 had done so sometimes. The study took approximately 15 minutes to complete, and each participant received a $5 USD gift card as an honorarium for their time.

3.2 Experimental Design and Interruption Interfaces

The study used a 2×2 between-subjects factorial design, manipulating two factors: *interface condition* and *VR task*.

In the *baseline* interface condition, a Google Pixel 3a mobile phone was mounted in a 3D-printed shell and acted like a doorbell. The phone ran an Android application with a high-contrast button labeled *"Tap Here To Seek VR User's Attention"*, which sent notifications like *"Someone is on your left!"* to the VR user. The static interface was positioned on a desk adjacent to the monitor displaying the VR user's activity (Figure 1b; Appendix C), ensuring it was visible from the participant's seated position in the study space throughout the session. This placement was chosen as the monitor displayed the real-time VR content and thus served as a natural point of visual interest. The location provided sufficient clearance for participants to approach and use the interface without entering the VR user's movement volume. While positioning the interface directly in front of the VR user could have been an alternative, as the VR user's orientation could shift during VR activities, their movements could have occluded the interface or made it unsafe for participants to approach.

[1] All participants self-identified their gender.

In the *proactive* interface condition, a proactive proxy interface anticipated the bystander's intent to interrupt and offered relevant interruption options. This interface featured a Google Pixel 3a housed in a 3D-printed shell mounted atop a custom-built robot (Figure 1c; Appendix C). The robot, resembling a teleconferencing device, stood 100 cm tall and was built from T-slotted extruded aluminum. A Raspberry Pi 4 controlled two Sparkfun dual motor drivers (i.e., TB6612FNG), which powered four 65×25 millimeter wheels via Adafruit DC motors (i.e., ID: 3777). The robot was operated using a Wizard-of-Oz approach by one of the experimenters (i.e., the wizard) to ensure reliable interception of participants' communication cues and consistent robot behavior across trials. The wizard monitored the study session through a webcam and controlled the robot's movement via a Flask-based web interface. The wizard initiated robot movement as soon as a participant explicitly expressed intent to communicate with the VR user, either by approaching the VR user and speaking to them or by calling out to them from a distance. The robot's starting position was identical to the location of the static interface in the baseline condition, and its height (100 cm) matched the height at which the static interface was positioned, ensuring equivalent initial visibility across conditions. The robot traveled in a straight path toward the participant at a slow walking pace (approximately 0.3–0.5 m/s) and stopped at a distance of 40–60 cm from the participant [48]. Both the speed and stopping distance were chosen to minimize the risk of startling participants or causing accidental collisions. Upon stopping, the robot played a brief notification sound (i.e., beep) to signal that it had completed its movement and was ready for interaction.

As a control, during both interface conditions, the HTC Knock Knock interface was displayed on the monitor connected to the HTC Vive Pro that the experimenter was wearing. This interface could also be used by participants if they desired.

Following O'Hagan et al. [50], the study included two VR applications to reflect real-world diversity in user engagement: i.e., an *active* task and a *passive* task. The active task involved playing Open Saber VR [13], an open-source clone of Beat Saber, where the VR user stood and used full-body movements to slash through incoming beats synchronized to music. In the passive task, the VR user remained seated and watched a music video using a customized music video player, employing minimal body movements to simulate navigation within the virtual environment. Both applications were audio-intensive and delivered sound through the HTC Vive Pro headset's integrated headphones, effectively isolating the VR user from external auditory stimuli.

Each interface condition was tested 40 times (i.e., 20 with the active VR task and 20 with the passive VR task) to ensure balanced representation across both VR task types.

3.3 Procedure

The study took place in a lab divided into two main areas: a *study space*, where the participant completed math problems on a computer, and a semi-public *shared space*, where a VR user engaged in a VR activity. Executing the deception required distinct experimenter roles and a sequence of phases, detailed in the following subsections.

3.3.1 Experimenter Roles. The study required three experimenters, each assuming a specific role during the session (Figure 2a).

Host: The host conducted the study and was the only experimenter that the participant recognized as part of the research team.

VR User: The VR user played a VR game or watched a VR video while wearing an HTC Vive Pro headset. Participants perceived this person as a fellow occupant of the shared space, not as someone affiliated with the study.

Wizard: The wizard operated behind the scenes, remotely monitoring the session via a Flask-based web interface that linked the VR applications, interruption interfaces, participant computer, and a web camera overlooking the study and shared spaces. Through this interface, the wizard could control the robot, trigger a lockout of the participant's computer using a C# script, and monitor safety in case of accidental physical contact between the participant and VR user (e.g., an unintended arm swing). No such incidents occurred during the study. The wizard remained hidden and was never introduced to participants.

3.3.2 Study Phases. The deception study progressed through four sequential phases, as illustrated in Figure 2b.

Phase 1: Introduction. The host welcomed the participant and invited them to sit in the *study space*. After sitting, the participant was given a consent form to read and sign and the host then explained that the study involved solving math problems for 15 minutes on the computer in front of the participant. When the participant was ready, a timer was started and the participant began the task.

Phase 2: Deception. During this phase, the participant quickly completed basic math problems on the computer. After solving math problems (approximately two minutes), the host informed the participant they needed to step out for an urgent matter. The host instructed the participant to continue working until the 15-minute session ended and assured them they would return within five minutes. Before leaving, the host pointed out an individual in the *shared space* and advised the participant to seek help from them if needed. Unbeknownst to the participant, this individual was another experimenter playing the role of a VR user. To avoid raising suspicion or prompting questions about interacting with someone in VR, the VR user was not yet wearing the VR headset. The VR user made eye contact with the participant and offered a friendly gesture such as a wave, smile, or nod, to signal their availability as a potential source of help.

Phase 3: Interruption. After the host left, the VR user put on the headset and began engaging with a VR game or video. The audio was loud enough to leak from the headset, making their activity audible to the participant. Once the VR session began, the wizard remotely locked the participant out of the study computer, prompting them to seek help from the VR user. During this time, participants had full autonomy: they could choose to interrupt the VR user, wait for the host to come back, or leave the study space. The session ended when the participant (i) interrupted the VR user, (ii) left the space, or (iii) waited more than two minutes (thus indicating a reluctance to interrupt).

Phase 4: Survey/Debriefing. After the study, participants completed a demographic and experiential survey (Appendix A) consisting of 8 Likert-scale items, 4 closed-ended questions, and 10

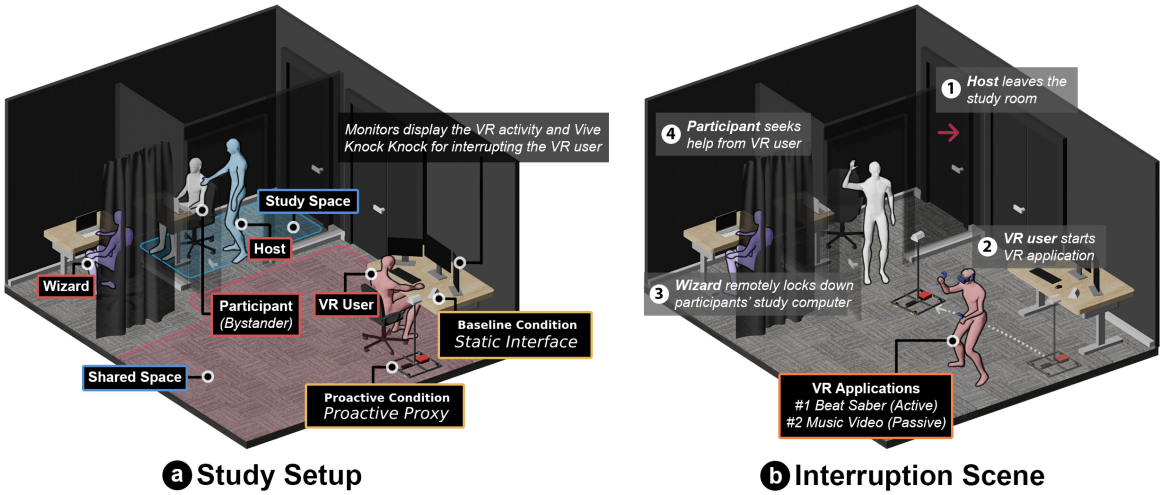

Figure 2: (a) The study was conducted within a room that was partitioned into three distinct spaces: a *study space*, where the participant was seated and completed math problems on a computer; a *shared space*, where a VR user engaged in a VR activity and interruption interfaces were available; and a *wizard area*, which was inaccessible to both the participant and VR user. (b) The experiment involved four phases: (1) the host exited the room, (2) the VR user began one of the two VR applications, (3) the wizard remotely locked the participant's computer after two minutes, and (4) the participant was compelled to seek assistance from the VR user.

open-ended questions about their experience interrupting the VR user. Although each participant had the opportunity to interact with only one interface (i.e., static or proactive proxy) due to the between-subjects design, the survey included video clips demonstrating both interfaces, as well as a third method: physically tapping the VR user on the shoulder. In each clip, the VR user engaged with the same application that participants had encountered in their study, encouraging them to reflect on their own experience. Participants then selected their preferred interruption method, first assuming the VR user was a stranger (as in the study), and then assuming they knew the user. Finally, the host debriefed participants on the true purpose of the study. Because deception was involved, participants were then asked to re-consent to the use of their data. No participants declined consent, so all data was retained.

3.3.3 Additional Considerations. To account for the complexity of ad-hoc interactions, we additionally controlled for key variables including social dynamics, technical familiarity, and situational context. For example, the VR user and bystander were strangers, the bystander had no prior knowledge of the interruption interfaces, and the urgency to interrupt emerged organically rather than through direct prompting. Additional controls were also implemented to minimize confounding variables. To account for gender dynamics, we balanced the genders of the bystander and VR user across conditions by rotating the VR user role among three experimenters (i.e., two male, one female).

Since the study took place on a university campus, some participants were familiar with certain experimenters. In these cases, we assigned the VR user role to maintain unfamiliarity between the participant and the VR user. To minimize external influence, only the experimenters and participant were present during the study. Although the lab was a semi-public space, access was restricted to avoid the confounding effects of additional bystanders, which could complicate interaction dynamics and the data analysis.

3.4 Analysis

To analyze bystanders' experiences during spontaneous interactions with unfamiliar VR users across the two interface conditions, a mixed methods approach was used.

3.4.1 Quantitative Analysis. Video recordings from all 80 participants were divided among four researchers, who analyzed them to extract the study outcomes. Because participants received no explicit instructions on how to interrupt the VR user, participants employed a range of interruption strategies, some extending beyond the provided interruption interface. Researchers thus tagged each video recording with its *interruption outcome*, noting whether the VR user was interrupted and, if so, what *interruption strategy* was used. Given participants' autonomy in choosing how to act, certain interruption strategies could occur unevenly across conditions or not at all, rendering traditional condition-based statistical comparisons potentially uninformative. We therefore adopted an observation-driven analytical approach, first establishing which strategies emerged across conditions before structuring subsequent analyses around these observed behaviors.

Interruption time was calculated as the interval between the participant noticing the computer lockout and them successfully gaining the VR user's attention (or the experiment timing out due to their inaction or them leaving the study room). Responses to Likert-scale items were analyzed to assess participants' *interruption experience*, i.e., perceived difficulty, discomfort, and confidence while interrupting, and perceptions of the interruption interface if used to interrupt the VR user. We also examined participants' *preferred interruption strategy* for different bystander interruption interfaces across the different VR tasks, considering both acquainted and unacquainted VR users.

While VR task (active vs. passive) was included to represent a range of typical VR usage scenarios, a preliminary analysis revealed no significant main effects of VR task or interactions with interface condition on interruption strategy distribution (log-linear analysis: all $p > 0.65$), interruption timing (Mann–Whitney U tests: all $p > 0.05$), or interruption-experience measures (2 × 2 ART ANOVAs: all $p > 0.05$). The only exception was a significant effect of VR task on preferred interruption interface for acquainted users; we report this result in Section 4.1.4. Because VR tasks did not meaningfully influence any of the behavioral or experiential outcomes, and to simplify the presentation of results and focus on our primary comparison between interface conditions, we aggregated across VR tasks in all subsequent analyses. Full condition- and task-level breakdowns are available in Appendix B.

3.4.2 Qualitative Analysis. To understand bystanders' experiences during spontaneous interactions with unfamiliar VR users, we conducted a thematic analysis of the responses to the 10 open-ended survey questions [5, 6]. A total of 531 responses were compiled into a spreadsheet, with columns representing each question and rows corresponding to individual participants. Three researchers independently reviewed an initial subset of 20% of these responses, using open-coding, loosely following Saldaña's first-cycle approach [60]. Through iterative discussions, the entire research team refined these initial codes into a codebook containing 12 codes. The codebook was then used to code the remaining responses. After individual coding was complete, the three researchers reconvened to review the data, cluster related concepts, and refine the codes into the final themes presented below. As our thematic analysis utilized all participant responses, the themes reflected experiences reported by participants in both VR tasks. For transparency, we indicate each participant's study configuration when presenting quotations.

4 Results

While our study took place in a shared lab space, the experimental design effectively recreated the urgency and realism of real-world interruptions. This was reflected in post-debriefing comments such as *"the only thought in my mind was to get the VR user's attention quickly since I was supposed to finish these math questions quickly"* (P20, **Condition:** Baseline/**Task:** Active) and *"it all felt very natural, I had no idea what was actually going on"* (P72, Proactive/Active). Herein, we present the results from the video data, Likert items, and preference questions, and then explore the themes that were identified within participants' open-ended responses.

4.1 Quantitative Results

We first explain the interruption outcomes and strategies that were observed, and then explore their influence on interruption time and interruption experience.

4.1.1 Interruption Outcomes and Strategies. Across the video analysis of all 80 participants, three distinct interruption outcomes were observed: **Physical Interruption**, i.e., participants physically interacted with the VR user to get their attention; **Proactive Proxy**, i.e., participants used the proactive proxy interface; and **Timeout**, i.e., participants did not interrupt the VR user at all, either leaving the experiment area or remaining inactive for more than two minutes.

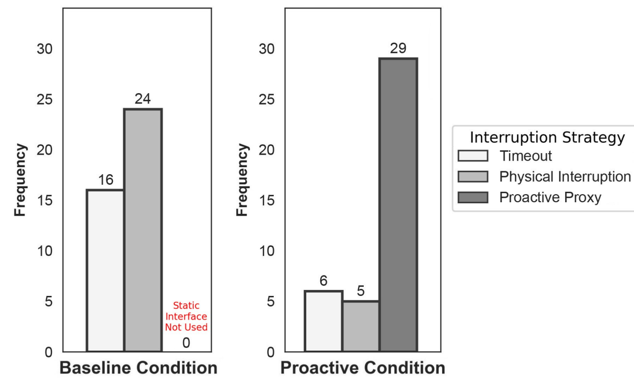

Figure 3: The interruption outcomes observed during the (a) baseline condition and (b) proactive condition. Note that in the baseline condition, none of the participants used the static interface, hence the "0" is reported.

In the baseline condition, 24 of the 40 participants (60%) chose to interrupt the VR user, all by physically interacting with them. The remaining 16 participants (40%) opted not to interrupt and instead waited for the host to return (Figure 3a). Although the static interface was clearly visible, none of the participants used it to interrupt, echoing prior findings on design blindness [43]. In the proactive condition, 34 participants interrupted the VR user (85%). Of these, 29 participants (72.5%) used the proactive proxy and 5 participants (12.5%) opted for physical interruption. Only 6 participants (15%) timed out without interrupting (Figure 3b). Among those who timed out, only 2 participants reported having failed to notice the proactive proxy.

To examine whether the distribution of interruption strategies differed between conditions, we conducted a 2 × 3 Fisher's Exact Test with Monte Carlo simulation (10,000 replicates). The test revealed a significant association between condition and strategy ($p < 0.001$), indicating that the distribution of interruption strategies differed across conditions. These findings suggest that interruption strategies incorporating design elements such as motion, proximity (within arm's reach), and attention can influence whether users engage with them to interrupt others. However, even with these features, some participants still chose not to use the available interruption methods.

As shown in Figure 3, participants in the baseline condition either physically interrupted the VR user or timed out. None used the static interface. Similarly, some participants in the proactive condition also opted for physical interruption or timed out rather than using the available proxy. The emergence of these alternative strategies, combined with the absence of static interface usage, meant that direct statistical comparisons between interface conditions would not yield representative insights into participants' experiences. Analyzing the results by observed strategy aligned more closely with what actually occurred and better captured how different interruption behaviors shaped participant experiences.

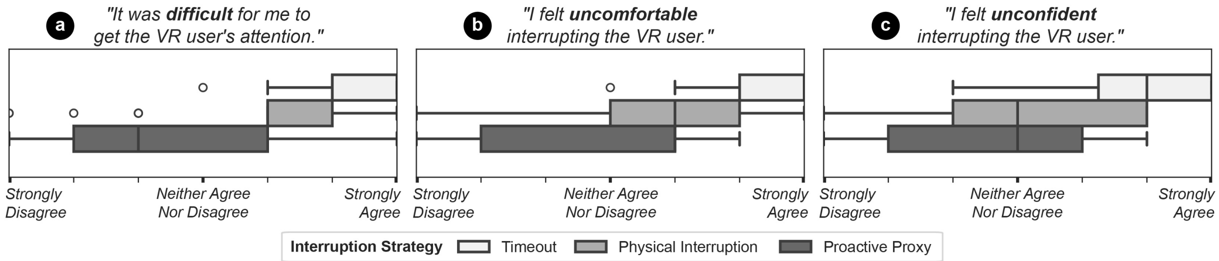

Figure 4: Participants' response distributions to the interruption experience Likert-scale items regarding the difficulty of getting the VR user's attention, how uncomfortable they were, and how unconfident they felt.

The following analyses of interruption time and experience considered the three observed strategies: physical interruption, proactive proxy use, and timeout. Where strategies occurred across both conditions (i.e., physical interruption and timeout), data was pooled due to the limited sample sizes within the individual conditions. Condition-level breakdowns for all measures are provided in Appendix B.

4.1.2 Interruption Time. We report the interruption times for the two interruption strategies that occurred (i.e., physical interruptions and interruptions using the proactive proxy). Since physical interruptions occurred during the baseline and proactive proxy conditions, the reported times for physical interruptions include participants from both conditions. Participants who used the proactive proxy (*Mean* = 28.5 seconds, *SD* = 9.2 seconds) were able to interrupt the VR user faster than those who physically interrupted the VR user (*Mean* = 35.1 seconds, *SD* = 15.4 seconds), with a mean difference of 6.6 seconds. This difference is likely due to the design of the proactive proxy, which moved toward the participant and eliminated the need for them to approach the VR user, thus streamlining the interruption process.

4.1.3 Interruption Experience. Overall, participants reported a moderately negative experience interrupting the VR user, which is unsurprising given the nature of the situation. Using ANOVA with the Aligned Rank Transform [83] revealed significant differences in participants' ratings of difficulty, discomfort, and confidence across interruption strategies (Figure 4). The analysis revealed a significant main effect of interruption strategy for each measure (i.e., difficulty: $F_{2,74} = 27.95$; $p < 0.001$; discomfort: $F_{2,74} = 23.35$; $p < 0.001$; confidence: $F_{2,74} = 17.28$; $p < 0.001$). Post-hoc comparisons using ART-C [15] with a Holm-Bonferroni correction confirmed that all three interruption strategies differed significantly from one another on each measure ($p < 0.001$ for all nine comparisons). The proactive proxy yielded the most favorable experience: participants found it easiest to use, felt most comfortable, and reported the highest confidence. Physical interruptions resulted in moderate discomfort and lower confidence, though participants found it slightly easier to get the VR user's attention. Timing out led to the most negative experience, with participants feeling uncomfortable, uncertain, and struggled to interrupt while waiting for the host to return.

Participants who used the proactive proxy interface ($N = 29$) additionally reported that it was easy to use, comfortable, and increased their confidence in attracting the VR user's attention (all ratings: $Mdn = 6.0$, $IQR = 1.0$). The interface was also perceived to be neither difficult to notice ($Mdn = 2.0$, $IQR = 1.0$) nor difficult to use ($Mdn = 1.0$, $IQR = 1.0$).

4.1.4 Preferred Interruption Strategies. Chi-Square Goodness-of-Fit tests were used to examine whether participants' preferences deviated from an equal distribution across the three interruption strategies. When asked which interruption strategy participants preferred to interrupt a stranger (Figure 5a), participants in the baseline condition showed a significant preference for the proactive proxy ($\chi^2(2, N = 40) = 13.38$, $p = 0.0012$), with 60% (24/40 participants) selecting it. The static interface was preferred by 25% (10/40) and physical interruption by 15% (6/40). Participants in the proactive condition showed a similar trend, favoring the proactive proxy (18/40) over the static (12/40) and physical strategies (10/40), but this preference was not statistically significant ($\chi^2(2, N = 40) = 2.6$, $p = 0.27$). These results suggest that participants in the baseline condition recognized the need for a more effective interruption strategy.

In contrast, when asked about interrupting a familiar VR user (Figure 5b), participants showed much greater preference for physical interruptions. Participants in the baseline condition opted for physical interruption 50% (20/40) of the time, with the static interface and proactive proxy each chosen by 25% (10/40), although this preference was not statistically significant ($\chi^2(2, N = 40) = 5.0$, $p = 0.08$). Participants in the proactive condition showed a strong preference for physical interruptions ($\chi^2(2, N = 40) = 15.6$, $p < 0.001$), with 62.5% (25/40) selecting them, compared to 15% (6/40) for the static interface and 22.5% (9/40) for the proactive proxy.

Because VR task was a potential factor shaping these preferences, we examined whether preferred interruption strategies differed between the active and passive VR-task groups. A 2 × 3 Fisher's Exact Test with Monte Carlo simulation (10,000 replicates) revealed a significant association between VR task and preferred interface when participants considered interrupting an acquainted user ($p = 0.012$). Participants in the passive-task condition showed a stronger preference for physical interruptions (29/40 physical, 6/40 static interface, 5/40 proactive proxy), whereas preferences in the active-task condition were more evenly distributed across the three strategies (16/40 physical, 10/40 static interface, 14/40 proactive proxy). While those who used the proactive proxy deemed it unnecessary for familiar users, the mixed preferences from those in the baseline condition suggest that even with familiar users, some form of interruption support may still be beneficial depending on the situation.

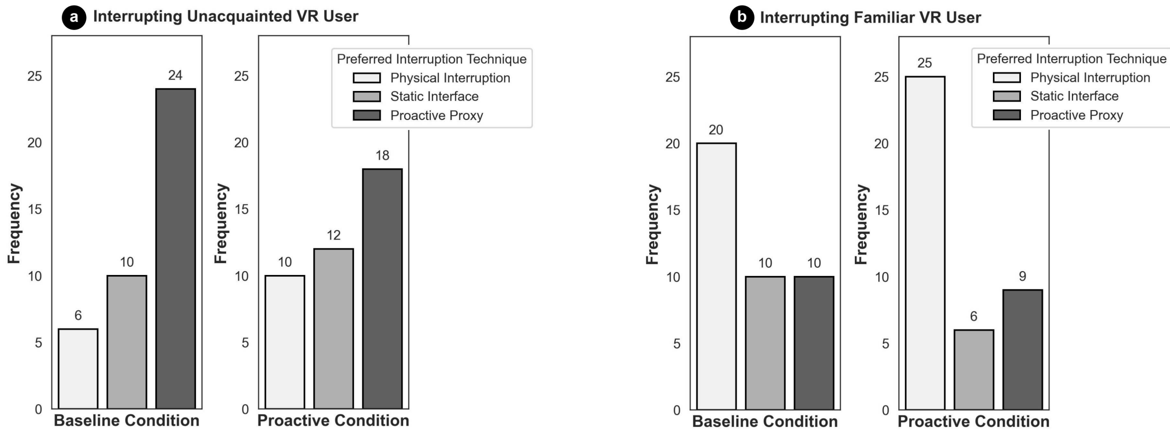

Figure 5: Participant preferences for the interruption strategy they would prefer to use to interrupt a VR user who is (a) a stranger and (b) familiar. (a)-left and (b)-left illustrate the preferences for participants who experienced the baseline condition, whereas (a)-right and (b)-right illustrates the preferences for participants who experienced the proactive condition.

4.2 Qualitative Results

Our analysis of the open-ended responses revealed concerns that negatively affected bystanders' experiences, highlighted the challenges posed by current static interruption interfaces, and identified the potential of proactive interruption interfaces to better support bystander engagement.

4.2.1 Concerns about Physical Interruptions. Participants typically treated physical touch as a last resort, first attempting verbal communication to gain the VR user's attention (e.g., *"I first planned on staying in my seat and just speaking to them from the other room. Once that didn't work, I got up and stood behind them and began speaking to them. After I saw that that was unsuccessful, I tapped them on the shoulder to get their attention"* (P29, Baseline/Passive)). While some found physical interruptions to be "*simple*", "*reliable*", and more "*human-like*", many expressed hesitation. This reluctance stemmed from three key concerns: (i) discomfort with physical contact, (ii) safety risks, and (iii) fear of startling the VR user.

Discomfort with Physical Touch. On several occasions, participants physically interrupted the VR user despite feeling discomfort. One participant noted, *"I am wary of physical touch with people I do not know since that feels like crossing a boundary without their permission"* (P26, Baseline/Passive), while another shared, *"it was not comfortable to interact with the VR user"* (P9, Baseline/Active). Both the Likert item ratings for physical interruption (Figure 4) and the interruption time reflect this discomfort. Interestingly, some participants employed more creative strategies such as lightly poking the VR user's arm with a pen or shaking their chair, as alternatives to direct physical touch.

Some participants were so uncomfortable with physical interruption that they opted not to interrupt at all. One participant shared, *"I tried to communicate verbally by saying hello, and then when that didn't work I moved closer to the VR user and tried to communicate verbally again ... I gave up because I felt uncomfortable getting even closer or physically tapping the person. I decided I'd rather do poorly on the math questions and potentially have to redo the study than disrupt the person"* (P28, Baseline/Passive). This highlights a tension between the desire to engage and the perceived inappropriateness of touching a stranger without consent. As noted earlier, participants showed greater preference for physical interruption methods when the VR user was someone they knew.

Safety. Participants also expressed concern about being accidentally struck by the VR user, particularly while playing Open Saber. P21 (Baseline/Active) shared feeling *"more nervous about [their] safety if [they] were to walk too close to the user or in front of them"*, and P19 (Baseline/Active) opted to *"approach from behind to be less in the way of the equipment"*. However, this concern was not limited to participants in the active condition. For instance, P48 (Proactive/Passive) remarked, *"I've seen many videos where VR users accidentally hit or run into others because they can't see"*, reflecting similar caution even during the passive VR task. These findings suggest that prior exposure to VR-related media may heighten awareness of potential safety risks when interacting with immersed VR users.

Fear of Startling the VR User. Although less common than concerns about discomfort or safety, some participants worried that physical interruption might startle the VR user. To avoid this, P18 (Baseline/Passive) opted for a *"slight nudge"* rather than a more forceful gesture, while P35 (Proactive/Active) waited until the headset music stopped, reasoning that the user would be less immersed and therefore less likely to be frightened: *"I just assumed it would be scary if I touched the user while they were immersed in a game"* (P35, Proactive/Active). These examples highlight how participants considered the VR user's state of immersion when choosing how and when to interrupt.

4.2.2 Static Interfaces Remain Unused Due to Poor Visibility and Privacy Concerns. Static interruption interfaces, which allow bystanders to alert the VR user via a button press, offer a potential solution to the challenges of physical interruption. These advantages resonated with participants during the post-study survey. For example, P38 (Baseline/Passive) remarked, *"I feel that it would be a lot better than trying to physically touch them or be loud enough* *for them to hear you. This way works better for me because I am not good at talking to others (I am very shy)"*, while P6 (Baseline/Active) noted, *"yeah absolutely I'd use that instead of accidentally getting whacked in the head by the VR user."*.

Despite this positive feedback, none of the participants who had access to the static interface chose to use it. As P7 (Baseline/Passive) noted, *"the desktop interface was unnoticeable for me because it was pretty far from my central focus (the VR user)"*. Several participants pointed out that such interfaces are *"not something that is commonly done"* (P37, Baseline/Passive) and highlighted the need for *"prior education that using the technique is the main way of interrupting the user"* (P1, Baseline/Active). These responses suggest that, contrary to assumptions held by VR headset manufacturers, use of static interruption methods cannot be taken for granted until such methods have increased noticeability and achieve broader normalization.

Beyond a general lack of awareness about the static interface, some participants deliberately avoided exploring the room, citing concerns about invading the VR user's privacy. As P7 (Baseline/Passive) explained, *"I do not like to tamper with people's stuff, so I did not really pay attention to the interface on the desk"*. The interface's placement further amplified these concerns, with P79 (Baseline/Passive) noting it was *"not very comfortable for me as it is too close to the VR user's personal belongings"*. Even when asked whether they would have used the interface had they noticed it, some participants remained hesitant. P27 (Baseline/Active) shared, *"I didn't notice it but I probably still wouldn't have used it if I had, this is because I don't want to touch other people's tech and risk doing something wrong or breaking something"*, while P10 (Baseline/Active) added, *"no, I would feel like I would be violating someone's privacy"*. These responses highlight the importance of designing interruption interfaces that are not only more visible but also clearly signal their purpose and boundaries to mitigate privacy-related discomfort.

4.2.3 Proactive Proxy Affordances. Participants who used the proactive proxy described it as "*easy*", "*safe*", and "*convenient*", with the Likert responses about interrupting the VR user reflecting a generally positive experience (Figure 4). Its mobility played a key role in making the interface noticeable and approachable. Many participants felt invited to engage, noting that the robot's movement signaled its purpose clearly. For instance, one participant remarked, *"it was very clear that the robot interface was for me since it seemed to move when I tried to get the VR user's attention"* (P36, Proactive/Active). Another shared, *"by blocking your way, it is obvious that you should interact with [the robot]"* (P57, Proactive/Active). P79 (Baseline/Passive), who was only exposed to the proactive proxy during the post-study survey, commented *"[With the robot interface], it feels like I am invited to communicate with the VR user"*. This sense of encouragement helped even hesitant participants feel more comfortable initiating interaction (e.g.,*"I felt uncomfortable because I didn't want to interrupt her, but the robot made me feel like it was okay to do so, and so when I pressed the button and she responded, I felt relieved"* (P42, Proactive/Active)).

The proactive proxy introduced a new dynamic: instead of participants initiating the interruption, they were first interrupted by the proxy itself. This shift was noted by participants, such as P67 (Proactive/Active) who shared, *"the motion of the robot really helped since the motion drew my attention to the robot first"*. Similarly, P30 (Proactive/Active) described how their initial plan to shout or tap the VR user changed when *"the robot approached me with the screen that said 'please tap to get the VR user's attention' ... I knew to tap that and felt like it would work"*. By actively engaging the participant, the proxy redirected their attention towards its interface, facilitating a more structured and reliable method of initiating contact.

Some participants raised concerns about privacy and practicality. Although the proactive proxy maintained a respectful distance (40-60 cm) to avoid collisions [48], this occasionally made its screen difficult to read, thus prompting participants to move closer. One participant remarked, *"I wasn't sure if the robotic touchscreen interface was meant for me to use or not. I did not want to touch someone else's property and I was too far away from the screen to read anything"* (P69, Proactive/Active). Others questioned the necessity of the system, describing it as *"over-engineered"* (P55, Proactive/Active) or *"overkill"* (P45, Proactive/Passive), while one noted, *"it may not be realistic to always have a robot that can follow you around"* (P72, Proactive/Active). A few participants also reported feeling *"startled"* (P39, Proactive/Active) or found the robot *"scary"* (P53, Proactive/Passive). These reactions suggest that public acceptance of proactive proxies may vary widely, shaped by factors such as aesthetics, perceived intrusiveness, and clarity of purpose.

5 Discussion and Future Work

Using a deception-based study methodology, we engineered authentic situational constraints that elicited genuine responses from bystanders. Our findings confirmed that bystanders experienced significant discomfort when interrupting unfamiliar VR users **(H1)**. This hesitation aligns with preliminary field observations by Wang et al. [76], which reported similar reluctance among bystanders to engage with VR users in public spaces. The static interface in our baseline condition was positioned in a natural, practical location for real-world deployment, i.e., adjacent to the monitor displaying real-time VR content and visible from the participant's seated position, yet it was overlooked by all participants **(H2)**. This attentional oversight is consistent with perceptual load theory [36] and prior work on public interactive systems, which showed that even prominently placed and clearly labeled interfaces can suffer from attentional blindness when passersby are cognitively focused on a task [32]. Interestingly, when the same interruption interface was mounted on a robotic proxy that proactively entered participants' field of attention, participants readily engaged with it and reported considerably less discomfort **(H3)**. Even if we consider the placement of the static interface to be a contributing factor to its lack of use, it raises questions about how precisely a VR user would need to anticipate bystander behavior for them to effectively position such an interface. By proactively presenting the interface, the proxy implicitly signaled that interrupting was socially appropriate. This supports prior findings that showed that people interrupt more sensitively when both motivation and contextual cues are present [11].

In the following sections, we examine the value of incorporating bystander perspectives in VR interaction design, explore key considerations for creating interfaces that support spontaneous encounters, and discuss the limitations of our research.

5.1 Need for Bystander-Centric Design in VR

The predominant approach to managing VR-bystander interactions has focused on systems that inform VR users about their physical surroundings. While recent work has shown that VR users' needs for such awareness vary by social context [52], the degree of immersion ultimately remains at the discretion of the VR user. This can create an imbalance in interaction control. If a user chooses full immersion, such as during a long-haul flight to avoid constant notifications about passenger movements, bystanders may have limited means to initiate contact.

Our findings demonstrate that bystander-centric interfaces can significantly improve bystanders' experience during impromptu encounters with fully immersed, unacquainted VR users. However, we do not propose replacing existing bystander awareness systems. These systems remain crucial for user safety, environmental awareness, and maintaining appropriate social boundaries in shared spaces. Instead, we envision them as one component of a broader solution that integrates the bystander's perspective and needs.

Drawing from our findings and previous research, we now explore how embedding bystander-centric design into VR systems can enhance the shared experience for both VR users and bystanders.

5.1.1 Communicating Availability. During our study, some participants chose not to interrupt the VR user after perceiving that they were deeply engaged. This behavior reflects Wang et al.'s [76] findings on in-the-wild VR interruptions and aligns with broader research on how people assess others' availability before initiating contact [67, 79, 80]. While prior work on availability management, ranging from physical cues like 'Do Not Disturb' signs to digital indicators of online status, showed that explicit signals make people more comfortable initiating interaction [80], VR lacks such mechanisms. Our study found that once bystanders noticed the existence of an interruption interface, they interpreted it as an implicit signal of availability. There thus appears to be a need for more practical methods for VR users to communicate their availability so bystanders can make more informed decisions.

5.1.2 Timing Interruptions Based on Urgency. In our experiment, participants could interrupt VR users by pressing a single button to trigger an interruption notification. However, current bystander awareness systems treat all bystander presence as equally urgent, lacking mechanisms to assess the priority of potential interactions [52]. To address this, bystander-centric interfaces should empower bystanders to control how their notifications are delivered, allowing urgent messages to be sent immediately while deferring lower-priority interruptions to natural transition points. This approach supports bystanders in expressing their communication needs without imposing unnecessary disruptions. At the same time, VR users can remain immersed in their tasks, responding to urgent matters promptly and engaging with less critical messages when appropriate.

5.2 Design Considerations for Bystander-Centric VR Systems

As one of the first controlled studies examining ad-hoc bystander-VR user interruptions in a naturalistic setting, we needed to anchor our investigation at a specific point along the continuum of possible encounters. As human interruptions vary along multiple dimensions (i.e., familiarity with the interruptee, degree of their immersion, urgency of the interruption, and social context), we chose the most challenging configuration: a fully immersed VR user who was a complete stranger, with bystanders having a moderate urgency to interrupt. This represented one end of the continuum. Interactions involving acquainted users, partially immersed users, or different urgency levels would likely present different barriers. Our deception-based methodology could thus serve as a template for future work that systematically varies these parameters, e.g., manipulating the VR user's degree of immersion, a bystander's relationship with them, the stakes of the interruption, or the presence of public onlookers.

Similarly, interruption interfaces may themselves exist on a continuum. We evaluated two extremes: a static interface representing current solutions that assume bystander awareness, and a proactive robotic proxy that actively engages bystanders upon detecting interruption intent. Our goal was not to propose robots as a practical everyday solution, but to isolate and demonstrate the benefits of proactive interface presentation. While we did not aim to identify an optimal solution, examining these extremes revealed valuable design insights for future bystander-centric VR systems. Notably, participants using the proactive proxy often assumed the static interface would suffice. This assumption did not hold when participants in the baseline condition experienced the limitations of this interface firsthand. These findings highlight the need for dedicated interruption interfaces and suggest a design continuum between passive and highly proactive approaches, with the ideal solution likely falling somewhere in between. From our observations, we identified three key design implications for supporting impromptu interactions between bystanders and unacquainted VR users in shared spaces:

D1. Minimize Physical Contact. Given participant concerns about discomfort, startle responses, and safety, future bystander-centric interfaces should prioritize non-contact interaction methods. While physical touch may be necessary in urgent cases, gesture recognition or interface elements placed outside the VR user's personal space would help minimize direct contact during routine interactions.

D2. Communicate Available Interactions. Since spontaneous encounters may involve bystanders unfamiliar with VR capabilities or available interruption methods, interfaces must proactively capture attention and clearly signal interaction possibilities. Drawing from public kiosk research [30], future designs should incorporate attention-directing cues such as motion, dynamic visuals, responsive lighting, or approach-triggered audio feedback to reduce interaction blindness. Clear labeling and intuitive affordances can further help bystanders understand how to initiate contact. In predictable settings, such as when sitting in adjacent airplane seats, interfaces mounted on a VR headset's side panels could display prompts like "Say 'Hi' to get my attention". As more position-constrained scenarios emerge, such context-aware interface designs warrant further exploration.

D3. Signal Social Permission. Our findings suggest that social permission must be signaled on two levels: permission to use the interface and permission to interrupt the VR user. For the former, future designs should clearly distinguish bystander interfaces from personal devices through visual styling, invitational labeling, and placement outside personal space. For the latter, interfaces could explicitly communicate the VR user's availability or willingness to be interrupted, empowering VR users to actively signal whether they wish to be disturbed rather than leaving bystanders to make assumptions. Together, these cues can reduce hesitation and encourage interaction in shared spaces.

5.3 Limitations

Our controlled methodology enabled the observation of authentic bystander behavior by fixing key parameters including the use of a single bystander, a fully immersed stranger, and an urgent need for the bystander to interrupt the immersed stranger. However, several additional factors should be explored to better understand bystander-VR user interactions and interruption interfaces.

5.3.1 Sample Demographics. Our university setting limited participants to primarily 18-30 year old students. Older adults, who may be less familiar with VR technology, could experience even greater discomfort approaching immersed strangers, while younger individuals might hold different expectations around technology-mediated interaction. While we balanced gender pairings, we did not examine gender effects on interaction dynamics. Cultural values also likely shaped interruption behaviors within our diverse sample, though these influences were not controlled. Furthermore, although our sample was culturally diverse, participants were situated within a North American university context. Norms around personal space, acceptable physical touch between strangers, and willingness to initiate unsolicited interaction vary substantially across cultures—for instance, acceptable interpersonal distances and attitudes toward interrupting unknown others differ between collectivist and individualist societies, and between high-contact and low-contact cultures [24, 26, 64]. The discomfort and hesitation we observed may therefore manifest differently in other cultural contexts. Future research should systematically explore these demographic and cultural factors.

5.3.2 Study Environment. Conducting our study in a controlled lab with a PC VR setup allowed for consistent experimentation, offering a foundational look at impromptu bystander-VR interactions. We isolated a single bystander to examine interface effects without confounding social dynamics, though in practice, the presence of other people may introduce social pressure, conformity effects, or fear of judgment that further shape interruption behavior. However, this setting lacked the complexity of real public space, where additional bystanders and varied environments (e.g., cafes, malls, or airplanes) can shape interaction dynamics. Real environments also introduce distractions (e.g., obstacles, noise, pedestrian flow) that could reduce interface visibility and further discourage interruption attempts. Factors such as the presence or absence of a monitor displaying the VR user's view or forced proximity between a bystander and VR user may also heighten or reduce interaction dynamics. Understanding how location affects these interactions remains an open question for future research.

5.3.3 Interruption Urgency. The urgency embedded in our study design likely shaped participant behavior. Notably, even with task completion motivating them, 40% of baseline participants avoided interrupting the VR user entirely. In lower-stakes scenarios, avoidance rates would likely be higher, and the barriers we observed would likely be amplified. This suggests our findings may represent a conservative estimate of bystander hesitation, highlighting the need for future research on how urgency influences bystander interventions.

5.3.4 VR Awareness Features. Our study used a PC VR setup with integrated headphones, creating full audiovisual immersion. Modern VR systems now include awareness features like Meta Quest's *Space Sense* or Apple Vision Pro's *People Awareness*, alongside open-ear audio and passthrough options. While our study focused on scenarios where users opt for full immersion, exploring how partial awareness shapes interactions could offer valuable insights. We see bystander-centric interfaces as complementary to these systems, and future research should examine how they can be integrated to create more holistic solutions.

Taken together, these factors suggest the barriers we observed may be amplified in real-world contexts, reinforcing the need for interfaces that proactively support bystander-VR user interactions.

6 Conclusion

As VR systems increasingly extend beyond private spaces, understanding and supporting interactions between VR users and bystanders becomes essential. Our research provides empirical evidence that bystanders often feel significant discomfort when interrupting unfamiliar VR users and that traditional static interruption interfaces frequently go unnoticed during spontaneous encounters. Our proactive proxy interface, on the other hand, demonstrated that explicitly discoverable interfaces can significantly improve bystander comfort and interaction quality by signaling social permission to initiate contact. These findings suggest that as VR systems increasingly move into shared and public environments, their design must also consider the needs of nearby bystanders. In particular, incorporating interfaces that clearly communicate availability and guide bystander attention can help trigger interruptions in a comfortable and socially appropriate way.

We advocate for integrating bystander-centric design into VR systems, demonstrating that giving bystanders agency in managing interruptions enhances comfort and engagement. These interfaces can complement existing awareness systems by offering explicit communication channels without compromising immersion. We outline key design considerations for such systems and suggest that future research explore how to effectively combine these approaches across diverse contexts. As VR adoption expands into offices, transit systems, and public venues, this integration will be vital for creating socially attuned VR experiences that respect the needs of all participants.

References

[1] Fouad Alallah, Ali Neshati, Yumiko Sakamoto, Khalad Hasan, Edward Lank, Andrea Bunt, and Pourang Irani. 2018. Performer vs. observer: whose comfort level should we consider when examining the social acceptability of input modalities for head-worn display?. In *Proceedings of the 24th ACM Symposium on Virtual*

Reality Software and Technology (Tokyo, Japan) *(VRST ’18)*. ACM, New York, NY, USA, Article 10, 9 pages. doi:10.1145/3281505.3281541

[2] American Libraries Magazine. 2018. In a Virtual World: Virtual Reality in Libraries. https://americanlibrariesmagazine.org/2018/03/01/virtual-world-virtual-reality-libraries/. Accessed: 2026-01-23.

[3] Apple. 2023. Introducing Apple Vision Pro: Apple’s First Spatial Computer. https://www.apple.com/newsroom/2023/06/introducing-apple-vision-pro/. Accessed: 2026-01-23.

[4] Apple Inc. 2024. Use Apple Vision Pro on an Airplane or Train with Travel Mode. https://support.apple.com/en-us/118510. Accessed: 2026-01-23.

[5] Virginia Braun and Victoria Clarke. 2019. Reflecting on Reflexive Thematic Analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (Aug. 2019), 589–597. doi:10.1080/2159676X.2019.1628806

[6] Virginia Braun and Victoria Clarke. 2021. One Size Fits All? What Counts as Quality Practice in (Reflexive) Thematic Analysis? *Qualitative Research in Psychology* 18, 3 (July 2021), 328–352. doi:10.1080/14780887.2020.1769238

[7] Yingfeng Chen, Feng Wu, Wei Shuai, Ningyang Wang, Rongya Chen, and Xiaoping Chen. 2015. KeJia Robot–An Attractive Shopping Mall Guider. In *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings* (Paris, France). Springer-Verlag, Berlin, Heidelberg, 145–154. doi:10.1007/978-3-319-25554-5_15

[8] Chuan-Han Cheng, Chih-Yen Chen, Jie-Jhong Liang, Ting-Nan Tsai, Chih-Yin Liu, and Tzuu-Hseng S. Li. 2017. Design and Implementation of Prototype Service Robot for Shopping in a Supermarket. In *Proceedings of the 2017 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*. IEEE, Taipei, Taiwan, 46–51. doi:10.1109/ARIS.2017.8297181

[9] Yi-Shiu Chiang, Ting-Sheng Chu, Chung Dial Lim, Tung-Yen Wu, Shih-Huan Tseng, and Li-Chen Fu. 2014. Personalizing Robot Behavior for Interruption in Social Human-Robot Interaction. In *2014 IEEE International Workshop on Advanced Robotics and Its Social Impacts (ARSO)*. IEEE, Evanston, IL, USA, 44–49. doi:10.1109/ARSO.2014.7020978

[10] Herbert H. Clark. 1996. *Using Language*. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511620539

[11] Laura Dabbish and Robert E. Kraut. 2004. Controlling interruptions: awareness displays and social motivation for coordination. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) *(CSCW ’04)*. ACM, New York, NY, USA, 182–191. doi:10.1145/1031607.1031638

[12] Bella M. DePaulo and Jeffrey D. Fisher. 1980. The Costs of Asking for Help. *Basic and Applied Social Psychology* 1, 1 (March 1980), 23–35. doi:10.1207/s15324834basp0101_3

[13] dhcdht. 2019. Open Saber VR. https://github.com/dhcdht/OpenSaberVR. GitHub repository. Accessed: 2026-01-23.

[14] Pouya Eghbali, Kaisa Väänänen, and Tero Jokela. 2019. Social acceptability of virtual reality in public spaces: experiential factors and design recommendations. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia* (Pisa, Italy) *(MUM ’19)*. ACM, New York, NY, USA, Article 28, 11 pages. doi:10.1145/3365610.3365647

[15] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST ’21)*. ACM, New York, NY, USA, 754–768. doi:10.1145/3472749.3474784

[16] Future Travel Experience. 2024. Lufthansa Partners with Meta to Offer Mixed Reality Headsets in New Allegris Business Class. https://www.futuretravelexperience.com/2024/05/lufthansa-partners-with-meta-to-offer-vr-headsets-in-new-allegris-business-class/. Accessed: 2026-01-23.

[17] 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* (San Diego, CA, USA) *(DIS ’19)*. ACM, New York, NY, USA, 497–510. doi:10.1145/3322276.3322363

[18] Ceenu George, An Ngo Tien, and Heinrich Hussmann. 2020. Seamless, Bi-directional Transitions along the Reality-Virtuality Continuum: A Conceptualization and Prototype Exploration. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Porto de Galinhas, Brazil, 412–424. doi:10.1109/ISMAR50242.2020.00067

[19] Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander, Jeff Hotnog, 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 (April 2018), 1447–1456. doi:10.1109/TVCG.2018.2793698

[20] Erving Goffman. 1966. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Free Press, New York.

[21] David Golchinfar, Daryoush Daniel Vaziri, Gunnar Stevens, and Dirk Schreiber. 2022. Let’s Go to the Mall: Investigating the Role of User Experience in Customers’ Intention to Use Social Robots in a Shopping Mall. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS ’22)*. ACM, New York, NY, USA, 377–386. doi:10.1145/3532106.3533490

[22] Matt Gottsacker, Nahal Norouzi, Kangsoo Kim, Gerd Bruder, and Greg Welch. 2021. Diegetic Representations for Seamless Cross-Reality Interruptions. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Bari, Italy, 310–319. doi:10.1109/ISMAR52148.2021.00047

[23] Matt Gottsacker, Raiffa Syamil, Pamela Wisniewski, Gerd Bruder, Carolina Cruz-Neira, and Gregory Welch. 2022. Exploring Cues and Signaling to Improve Cross-Reality Interruptions. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Singapore, Singapore, 827–832. doi:10.1109/ISMAR-Adjunct57072.2022.00179

[24] Edward T. Hall. 1966. *The Hidden Dimension*. Doubleday, Garden City, NY.

[25] Guy Hoffman and Wendy Ju. 2014. Designing Robots with Movement in Mind. *J. Hum.-Robot Interact.* 3, 1 (feb 2014), 91–122. doi:10.5898/JHRI.3.1.Hoffman

[26] Geert Hofstede. 2001. *Culture’s Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations* (2nd ed.). SAGE Publications, Thousand Oaks, CA.

[27] Steven Houben and Christian Weichel. 2013. Overcoming interaction blindness through curiosity objects. In *CHI ’13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA ’13)*. ACM, New York, NY, USA, 1539–1544. doi:10.1145/2468356.2468631

[28] Ross Johnstone, Neil McDonnell, and Julie R. Williamson. 2022. When Virtuality Surpasses Reality: Possible Futures of Ubiquitous XR. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA ’22)*. ACM, New York, NY, USA, Article 6, 8 pages. doi:10.1145/3491101.3516396

[29] Brennan Jones, Yaying Zhang, Priscilla N. Y. Wong, and Sean Rintel. 2020. VROOM: Virtual Robot Overlay for Online Meetings. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA ’20)*. ACM, New York, NY, USA, 1–10. doi:10.1145/3334480.3382820

[30] Wendy Ju and David Sirkin. 2010. Animate Objects: How Physical Motion Encourages Public Interaction. In *Persuasive Technology*, Thomas Ploug, Per Hasle, and Harri Oinas-Kukkonen (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 40–51.

[31] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. 2010. A Communication Robot in a Shopping Mall. *IEEE Transactions on Robotics* 26, 5 (2010), 897–913. doi:10.1109/TRO.2010.2062550

[32] Jonas Keppel, Marvin Strauss, Sarah Faltaous, Jonathan Liebers, Roman Heger, Uwe Gruenefeld, and Stefan Schneegass. 2023. Don’t Forget to Disinfect: Understanding Technology-Supported Hand Disinfection Stations. *ACM Transactions on Computer-Human Interaction* 7, MHCI, Article 204 (Sept. 2023), 18 pages. doi:10.1145/3604251

[33] Sara Khan and Claudio Germak. 2018. Reframing HRI design opportunities for social robots: Lessons learnt from a service robotics case study approach using UX for HRI. *Future internet* 10, 10 (2018), 101.

[34] Ellyssa Kroski. 2021. *32 Virtual, Augmented, and Mixed Reality Programs for Libraries*. ALA Editions, Chicago.

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

[36] Nilli Lavie and Yehoshua Tsal. 1994. Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & psychophysics* 56 (1994), 183–197.

[37] Jingyi Li, Hyerim Park, Robin Welsch, Sven Mayer, and Andreas Butz. 2023. SeatmateVR: Proxemic Cues for Close Bystander-Awareness in Virtual Reality. *ACM Transactions on Computer-Human Interaction* 7, ISS, Article 438 (Nov. 2023), 20 pages. doi:10.1145/3626474

[38] Diana Löffler, Nina Schmidt, and Robert Tscharn. 2018. Multimodal Expression of Artificial Emotion in Social Robots Using Color, Motion and Sound. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) *(HRI ’18)*. ACM, New York, NY, USA, 334–343. doi:10.1145/3171221.3171261

[39] Daniel C. McFarlane. 1997. *Interruption of People in Human-Computer Interaction: A General Unifying Definition of Human Interruption and Taxonomy:*. Technical Report. Defense Technical Information Center, Fort Belvoir, VA. doi:10.21236/ADA333587

[40] 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* (Seoul, Republic of Korea) *(CHI ’15)*. ACM, New York, NY, USA, 2143–2152. doi:10.1145/2702123.2702382

[41] Meta. 2024. Getting Started With Travel Mode on Meta Quest 2 & 3. https://www.meta.com/blog/travel-mode-getting-started/. Accessed: 2026-01-23.

[42] Meta Quest Blog. 2021. A Smarter Guardian, Android Phone Notifications, and Improved Voice Commands in Latest Oculus Software Update. https://www.oculus.com/blog/a-smarter-guardian-android-phone-notifications-and-improved-voice-commands-in-latest-oculus-software-update/. Accessed: 2026-01-23.

[43] Jörg Müller, Dennis Wilmsmann, Juliane Exeler, Markus Buzeck, Albrecht Schmidt, Tim Jay, and Antonio Krüger. 2009. Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In *Pervasive Computing (Lecture Notes in Computer Science)*, Hideyuki Tokuda, Michael Beigl, Adrian Friday, A. J. Bernheim Brush, and Yoshito Tobe (Eds.). Springer, Berlin, Heidelberg, 1–8. doi:10.1007/978-3-642-01516-8_1

[44] Kristine S. Nagel, James M. Hudson, and Gregory D. Abowd. 2004. Predictors of availability in home life context-mediated communication. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) *(CSCW '04)*. ACM, New York, NY, USA, 497–506. doi:10.1145/1031607.1031689

[45] Carman Neustaedter, Samarth Singhal, Rui Pan, Yasamin Heshmat, Azadeh Forghani, and John Tang. 2018. From Being There to Watching: Shared and Dedicated Telepresence Robot Usage at Academic Conferences. *ACM Transactions on Computer-Human Interaction* 25, 6, Article 33 (Dec. 2018), 39 pages. doi:10.1145/3243213

[46] Cal Newport. 2016. *Deep Work: Rules for Focused Success in a Distracted World*. Grand Central Publishing, New York.

[47] Sara Nielsen, Mikael B. Skov, Karl Damkjær Hansen, and Aleksandra Kaszowska. 2023. Using User-Generated YouTube Videos to Understand Unguided Interactions with Robots in Public Places. *J. Hum.-Robot Interact.* 12, 1, Article 5 (Feb. 2023), 40 pages. doi:10.1145/3550280

[48] Mohammad Obaid, Eduardo B. Sandoval, Jakub Złotowski, Elena Moltchanova, Christina A. Basedow, and Christoph Bartneck. 2016. Stop! That is close enough. How body postures influence human-robot proximity. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, New York, NY, USA, 354–361. doi:10.1109/ROMAN.2016.7745155

[49] Joseph O'Hagan and Julie R. Williamson. 2020. Reality aware VR headsets. In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays* (Manchester, United Kingdom) *(PerDis '20)*. ACM, New York, NY, USA, 9–17. doi:10.1145/3393712.3395334

[50] 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* (Manchester, United Kingdom) *(PerDis '20)*. ACM, New York, NY, USA, 19–27. doi:10.1145/3393712.3395339

[51] Joseph O'Hagan, Julie R. Williamson, Mohamed Khamis, and Mark McGill. 2022. Exploring Manipulating In-VR Audio To Facilitate Verbal Interactions Between VR Users And Bystanders. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (Frascati, Rome, Italy) *(AVI '22)*. ACM, New York, NY, USA, Article 35, 9 pages. doi:10.1145/3531073.3531079

[52] Joseph O'Hagan, Julie R. Williamson, Florian Mathis, Mohamed Khamis, and Mark McGill. 2023. Re-Evaluating VR User Awareness Needs During Bystander Interactions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. ACM, New York, NY, USA, Article 876, 17 pages. doi:10.1145/3544548.3581018

[53] Joseph O'Hagan, Julie R. Williamson, Mark McGill, and Mohamed Khamis. 2021. Safety, Power Imbalances, Ethics and Proxy Sex: Surveying In-The-Wild Interactions Between VR Users and Bystanders. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Bari, Italy, 211–220. doi:10.1109/ISMAR52148.2021.00036

[54] Zhenhui Peng, Yunhwan Kwon, Jiaan Lu, Ziming Wu, and Xiaojuan Ma. 2019. Design and Evaluation of Service Robot's Proactivity in Decision-Making Support Process. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI '19)*. ACM, New York, NY, USA, 1–13. doi:10.1145/3290605.3300328

[55] Jenni Radun, Iida-Kaisa Tervahartiala, Ville Kontinen, Jukka Keränen, and Valtteri Hongisto. 2024. Do active noise-cancelling headphones' influence performance, stress, or experience in office context? *Building and Environment* 266 (2024), 112102. doi:10.1016/j.buildenv.2024.112102

[56] Maximilian Rettinger, Christoph Schmaderer, and Gerhard Rigoll. 2022. Do You Notice Me? How Bystanders Affect the Cognitive Load in Virtual Reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Christchurch, New Zealand, 77–82. doi:10.1109/VR51125.2022.00025

[57] Philip Riley. 1976. *Discursive and Communicative Functions of Non-Verbal Communication.* Technical Report. C.R.A.P.E.L., Universite de Nancy II.

[58] 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* (Barcelona, Spain) *(CHI PLAY '19)*. ACM, New York, NY, USA, 199–211. doi:10.1145/3311350.3347190

[59] Selma Šabanović, Sarah M Reeder, and Bobak Kechavarzi. 2014. Designing robots in the wild: In situ prototype evaluation for a break management robot. *Journal of Human-Robot Interaction* 3, 1 (2014), 70–88.

[60] Johnny Saldaña. 2016. *The Coding Manual for Qualitative Researchers* (3rd ed.). SAGE Publications, Thousand Oaks, CA.

[61] Paul Saulnier, Ehud Sharlin, and Saul Greenberg. 2011. Exploring minimal nonverbal interruption in HRI. In *2011 RO-MAN*. IEEE, Atlanta, GA, USA, 79–86. doi:10.1109/ROMAN.2011.6005257

[62] Valentin Schwind, Jens Reinhardt, Rufat Rzayev, Niels Henze, and Katrin Wolf. 2018. Virtual Reality on the Go? A Study on Social Acceptance of VR Glasses. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Barcelona, Spain) *(MobileHCI '18)*. ACM, New York, NY, USA, 111–118. doi:10.1145/3236112.3236127

[63] skidniks. 2024. Can Others Hear What You're Listening to given the Speakers Seem to Be Located on the Outer Part of the Strap? https://www.reddit.com/r/VisionPro/comments/1adjowl/can_others_hear_what_youre_listening_to_given_the/. r/VisionPro. Accessed: 2026-01-23.

[64] Agnieszka Sorokowska, Piotr Sorokowski, Peter Hilpert, Katarzyna Cantarero, Tomasz Frackowiak, Khodabakhsh Ahmadi, Ahmad M. Alghraibeh, Richmond Aryeetey, Anna Bertoni, Karim Bettache, et al. 2017. Preferred interpersonal distances: A global comparison. *Journal of Cross-Cultural Psychology* 48, 4 (2017), 577–592.

[65] Ryo Suzuki, Hooman Hedayati, Clement Zheng, James L. Bohn, Daniel Szafir, Ellen Yi-Luen Do, Mark D. Gross, and Daniel Leithinger. 2020. RoomShift: Room-scale Dynamic Haptics for VR with Furniture-moving Swarm Robots. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–11. doi:10.1145/3313831.3376523

[66] Ryo Suzuki, Eyal Ofek, Mike Sinclair, Daniel Leithinger, and Mar Gonzalez-Franco. 2021. HapticBots: Distributed Encountered-type Haptics for VR with Multiple Shape-changing Mobile Robots. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. ACM, New York, NY, USA, 1269–1281. doi:10.1145/3472749.3474821

[67] John C. Tang, Ellen A. Isaacs, and Monica Rua. 1994. Supporting distributed groups with a Montage of lightweight interactions. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work* (Chapel Hill, North Carolina, USA) *(CSCW '94)*. ACM, New York, NY, USA, 23–34. doi:10.1145/192844.192861

[68] thyazide. 2018. Remote Knock Knock. https://www.reddit.com/r/Vive/comments/8r8k85/remote_knock_knock/. r/Vive. Accessed: 2026-01-23.

[69] Meg Tonkin, Jonathan Vitale, Sarita Herse, Mary-Anne Williams, William Judge, and Xun Wang. 2018. Design Methodology for the UX of HRI: A Field Study of a Commercial Social Robot at an Airport. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) *(HRI '18)*. ACM, New York, NY, USA, 407–415. doi:10.1145/3171221.3171270

[70] Margaret V Tonkin. 2021. *Socially responsible design for social robots in public spaces.* Ph. D. Dissertation. University of Technology Sydney.

[71] UploadVR. 2016. HTC Vive "Knock Knock" Feature Lets You Know People in the Real World Still Care About You. https://www.uploadvr.com/new-knock-knock-feature-lets-you-know-someone-needs-attention/. Accessed: 2026-01-23.

[72] Maurizio Vergari, Tanja Kojić, Francesco Vona, Franca Garzotto, Sebastian Möller, and Jan-Niklas Voigt-Antons. 2021. Influence of Interactivity and Social Environments on User Experience and Social Acceptability in Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisbon, Portugal, 695–704. doi:10.1109/VR50410.2021.00096

[73] 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* (San Diego, CA, USA) *(DIS '19)*. ACM, New York, NY, USA, 487–496. doi:10.1145/3322276.3322334

[74] Emanuel Vonach, Clemens Gatterer, and Hannes Kaufmann. 2017. VRRobot: Robot actuated props in an infinite virtual environment. In *2017 IEEE Virtual Reality (VR)*. IEEE, Los Angeles, CA, USA, 74–83. doi:10.1109/VR.2017.7892233

[75] Yuntao Wang, Zichao (Tyson) Chen, Hanchuan Li, Zhengyi Cao, Huiyi Luo, Tengxiang Zhang, Ke Ou, John Raiti, Chun Yu, Shwetak Patel, and Yuanchun Shi. 2020. MoveVR: Enabling Multiform Force Feedback in Virtual Reality using Household Cleaning Robot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. ACM, New York, NY, USA, 1–12. doi:10.1145/3313831.3376286

[76] Yu Wang, Raphael Johannes Schimmerl, Martin Kocur, and Philipp Wintersberger. 2023. Ubiquity of VR: Towards Investigating Ways of Interrupting VR Users to Obtain Their Attention in Public Spaces. In *Virtual Reality and Mixed Reality*, Gabriel Zachmann, Krzysztof Walczak, Omar A. Niamut, Kyle Johnsen, Wolfgang Stuerzlinger, Mariano Alcañiz-Raya, Greg Welch, and Patrick Bourdot (Eds.). Springer Nature Switzerland, Cham, 40–52. doi:10.1007/978-3-031-48495-7_3

[77] Nethmini Thilakshi Weerawarna, Udaka Manawadu, and P. Ravindra S. De Silva. 2023. Sociable Robot 'Lometh': Exploring Interactive Regions of a Product-Promoting Robot in a Supermarket. *Journal of ICT Research and Applications* 17, 1 (28 Apr 2023), 58–81. doi:10.5614/itbj.ict.res.appl.2023.17.1.5

[78] Astrid Weiss, Nicole Mirnig, Ulrike Bruckenberger, Ewald Strasser, Manfred Tscheligi, Barbara Kühnlenz, Dirk Wollherr, and Bartlomiej Stanczyk. 2015. The interactive urban robot: user-centered development and final field trial of a direction requesting robot. *Paladyn, Journal of Behavioral Robotics* 6, 1 (2015), 000010151520150005.

[79] 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* (Boston, Massachusetts, USA) *(CHI '94)*. ACM, New York, NY, USA, 131–137.

doi:10.1145/191666.191726
[80] Mikael Wiberg and Steve Whittaker. 2005. Managing availability: Supporting lightweight negotiations to handle interruptions. *ACM Transactions on Computer-Human Interaction* 12, 4 (Dec. 2005), 356–387. doi:10.1145/1121112.1121114
[81] Julie R. Williamson, Mark McGill, and Khari Outram. 2019. PlaneVR: Social Acceptability of Virtual Reality for Aeroplane Passengers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI '19)*. ACM, New York, NY, USA, 1–14. doi:10.1145/3290605.3300310
[82] Jochen Wirtz, Paul G Patterson, Werner H Kunz, Thorsten Gruber, Vinh Nhat Lu, Stefanie Paluch, and Antje Martins. 2018. Brave new world: service robots in the frontline. *Journal of Service Management* 29, 5 (2018), 907–931.
[83] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 143–146. doi:10.1145/1978942.1978963
[84] 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* (Montreal QC, Canada) *(CHI EA '18)*. ACM, New York, NY, USA, 1–6. doi:10.1145/3170427.3188505

Appendix A: Survey Questionnaire

Note that Sections 1, 2, 3, and 6 were completed by all participants, whereas Sections 4 and 5 depended on what the participant did during the study (i.e., participants who did not use the provided interface completed Section 4, all other participants completed Section 5). Additionally, some phrasing and figure visualizations were slightly modified for each study configuration. Open-ended questions are indicated by an asterisk (*).

Section 1: Participant Demography

(1) Age
(2) Gender
(Male/Female/Non-Binary/Prefer not to say)
(3) How often do you use VR?
(5-point Likert Scale; 1: Never – 5: Always)
(4) How often have you felt the need to interact with a VR user?
(5-point Likert Scale; 1: Never – 5: Always)

Section 2: Overall Interaction Experience

The following questions contain 7-point Likert scales (1: Strongly Disagree; 7: Strongly Agree). In these questions, there will be a statement: please respond by circling the answer you think best describes what you think.

(1) It was **difficult** for me to get the VR user's attention.
(2) I felt **uncomfortable** interrupting the VR user.
(3) I felt **unconfident** interrupting the VR user.

Answer the next few questions descriptively to the best of your abilities.

(4) How did you plan on interrupting the VR user? Please describe your experience of communicating with the VR user after the PC crashed.*
(5) Would you like to give any other comments regarding your interaction experience?*

Section 3: VR Communication Interface

[Provided Interface: Static Interface**]** *In this study, the VR user had deployed a touchscreen interface on their desk for you to communicate with them. With this interface, instead of physically interrupting the VR user, you can tap the button on the screen to get their attention. (Figure A1)*

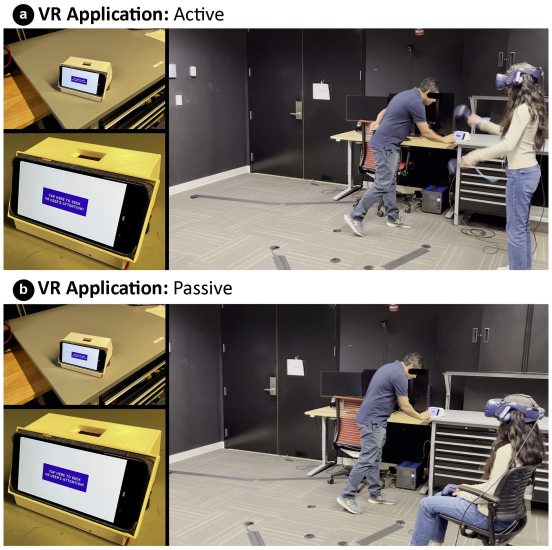

Figure A1: Screenshots from the video shown to participants during the survey who were assigned to the *Static Interface* condition, introducing them to the provided interface.

[Provided Interface: Proactive Proxy**]** *In this study, the VR user had deployed a robotic touchscreen interface for you to communicate with them. With this interface, the robot brings the controls to you so that you can interrupt the VR user from a distance. (Figure A2)*

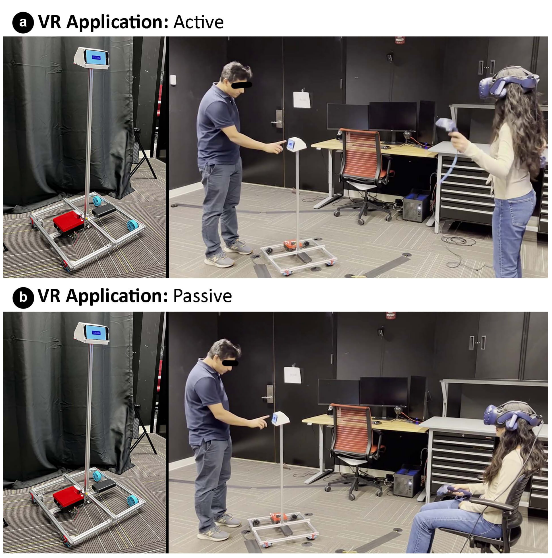

Figure A2: Screenshots from the video shown to participants during the survey who were assigned to the *Proactive Proxy* condition, introducing them to the provided interface.

(1) How do you feel about using this technique to interrupt the VR user?*
(2) Did you use this interface to communicate with the VR user?
(Yes/No)

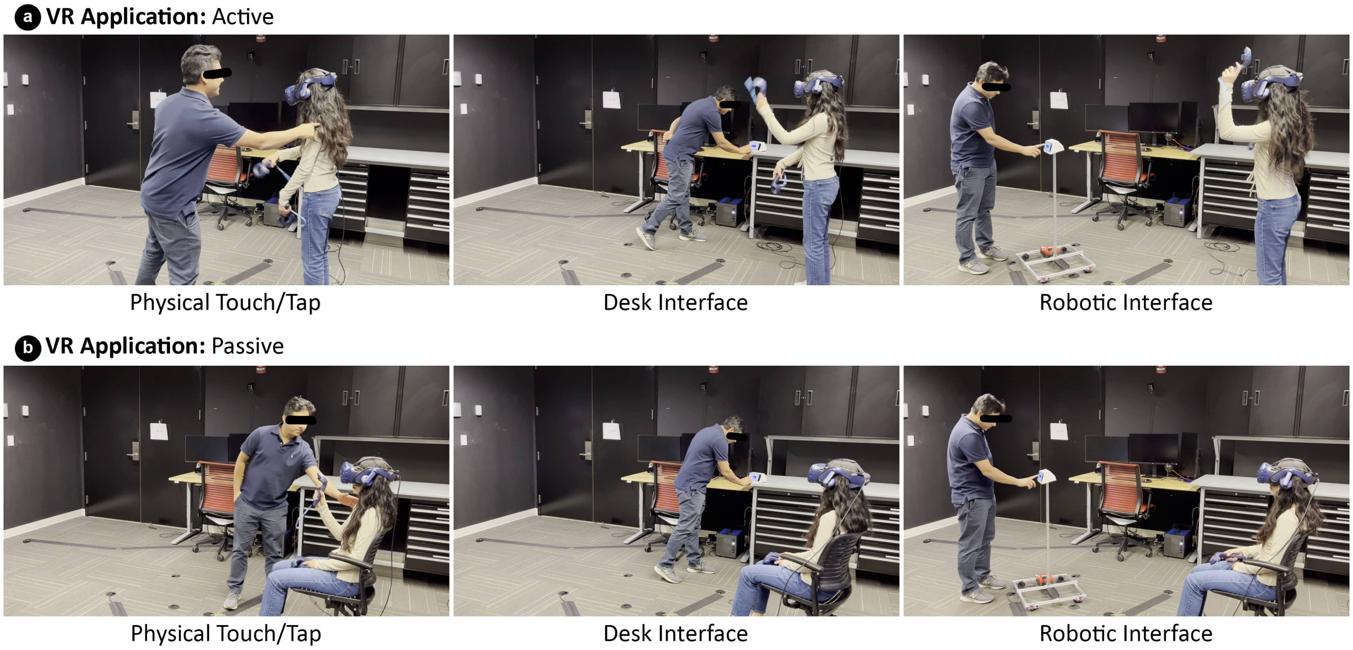

Figure A3: Screenshots from the video clips shown to participants who were assigned to the conditions (a) VR Application: *Active* and (b) VR Application: *Passive*, asking them about their interruption technique preference.

Section 4: Did Not Use Interface

(1) Why didn't you use this interface?
  (a) I did not notice the interface
  (b) I noticed the interface, but had other concerns

[If the interface was unnoticed]

(2) If you had noticed this interface, would it change your approach of communicating with the VR user? Please elaborate.*
(3) In your opinion, how do you think this interface can be made more noticeable?*

[If there were other concerns]

(4) Please elaborate your reason(s) for not using this interface to communicate with the VR user.*
(5) In your opinion, what changes to the interface would make you more willing to use it?*

Section 5: Interface Interaction Experience

The following questions contain 7-point Likert scale questions (1: Strongly Disagree; 7: Strongly Agree). In these questions, there will be a statement: please respond by circling the answer you think best describes what you think.

(1) It was **difficult for me to notice** the interface for communicating with the VR user.
(2) It was **difficult for me to use** the interface for communicating with the VR user.
(3) Having the interface **made it easier** for me to get the VR user's attention.
(4) Having the interface **made it comfortable** for me to get the VR user's attention.
(5) Having the interface **made me feel confident** interrupting the VR user.

Answer the next question descriptively to the best of your abilities.

(6) Would you like to give any other comments regarding your experience using this interface?*

Section 6: Interface Preference

[VR Task: Active**]** *Consider this scenario as you experienced in the study - the VR user is actively playing a game in VR. You need to interact with them. However, the music is loud, so they can't hear you. (Figure A3a)*
[VR Task: Passive**]** *Consider this scenario as you experienced in the study - the VR user is sitting on a chair and watching some content in VR. You need to interact with them. However, the music is loud, so they can't hear you. (Figure A3b)*

(1) Which of the three techniques - **physical touch/tap**, **a desk interface**, and **a robotic interface** - would you prefer to interrupt the VR user in the above situation? Please elaborate your choice.*
(2) For the same situation, if the VR user was someone you knew (friend/family/acquaintance), which technique would you prefer for interrupting? Please elaborate your choice.*

Appendix B: Configuration-Wise Measure Breakdown

To complement the aggregated findings reported in the main text, this appendix section provides configuration-wise breakdowns of all measures. These additional visualizations offer further context to support the transparency and reproducibility of our results.

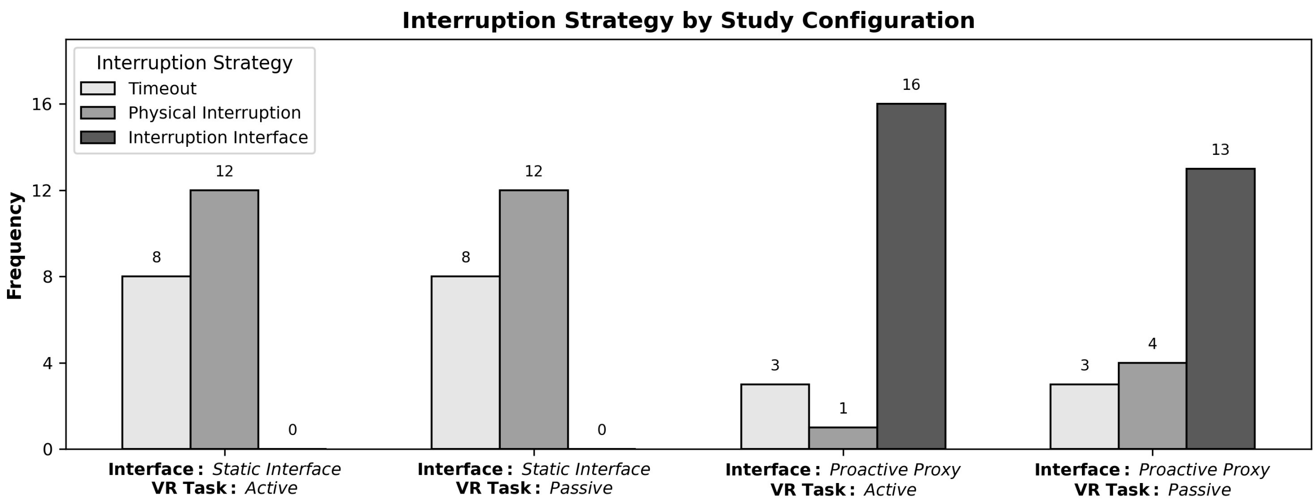

Figure B1: Interruption strategies by study configuration.

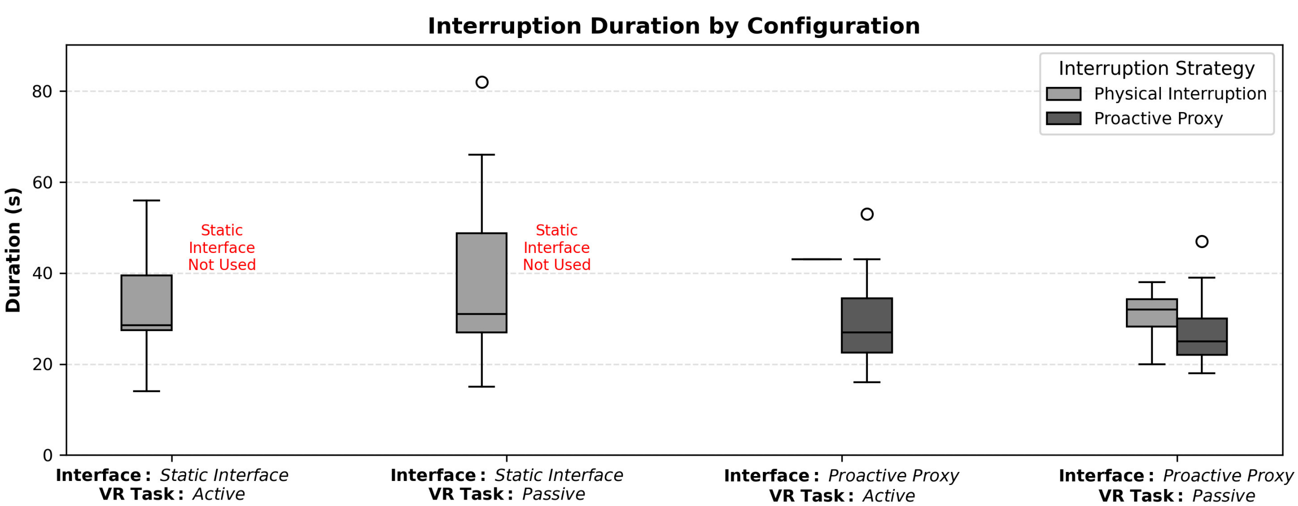

Figure B2: Interruption duration by study configuration and outcome. Note that there is no duration presented for timeout, as the interruption did not occur in that scenario.

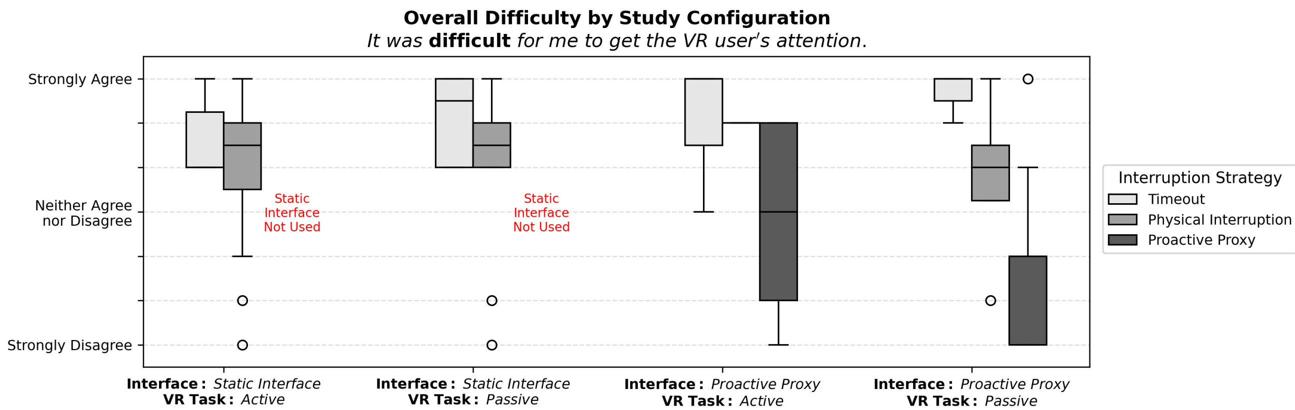

Figure B3: Perceived difficulty of getting the VR user's attention by study configuration and interruption strategy.

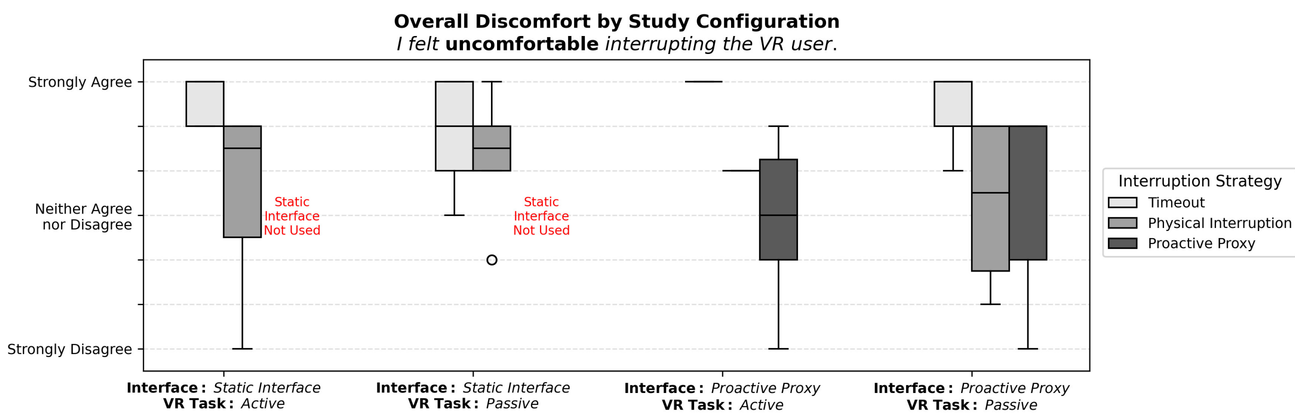

Figure B4: Perceived discomfort when interrupting the VR user by study configuration and interruption strategy.

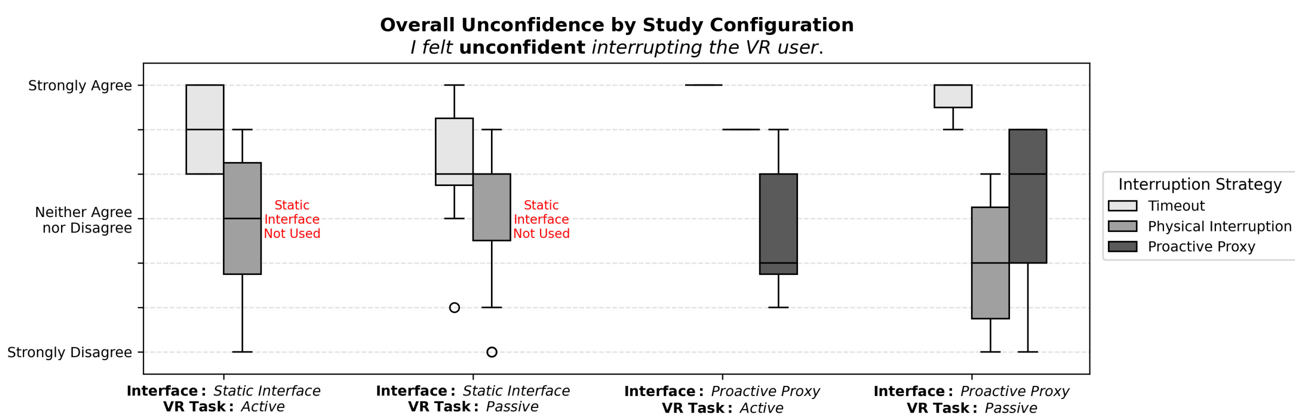

Figure B5: Perceived (un)confidence when interrupting the VR user by study configuration and interruption strategy.

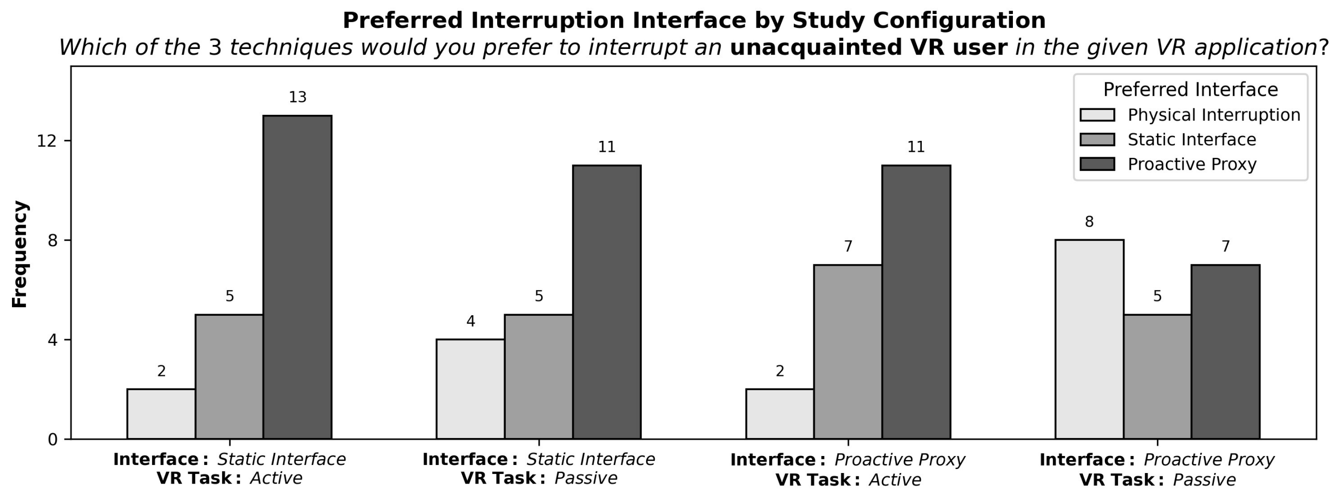

Figure B6: Preferred interruption interface by study configuration. After experiencing a single assigned configuration, participants were asked which of the three techniques (physical touch/tap, desk interface, or robotic interface) they would prefer to use to interrupt the VR user performing the same VR task they experienced in their study.

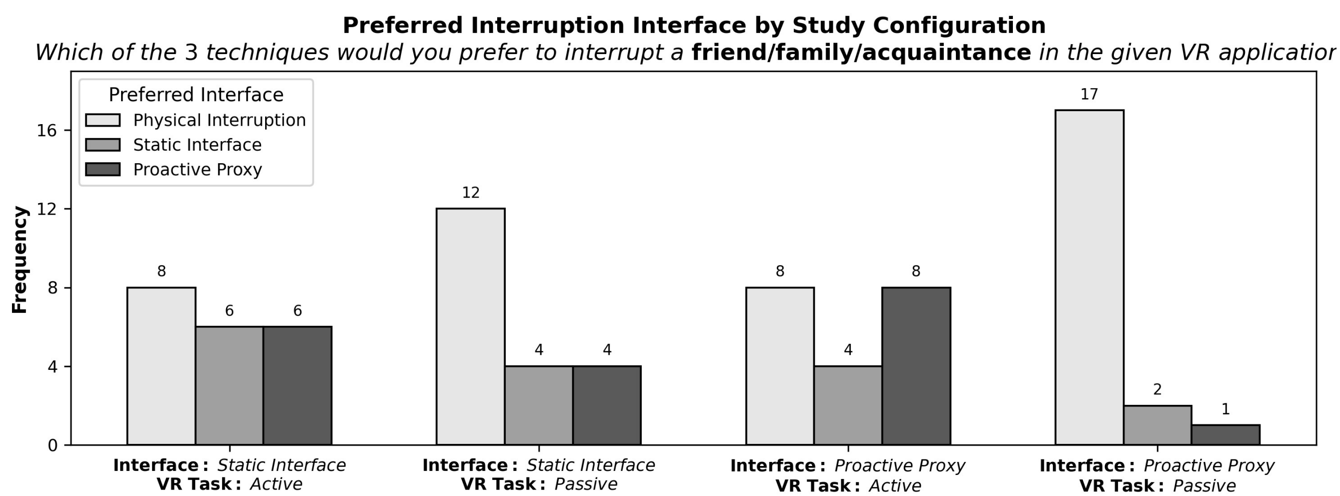

Figure B7: Preferred interruption interface by study configuration when interrupting someone they know. Participants indicated which technique they would prefer to use to interrupt a friend, family member, or acquaintance performing the same VR task they experienced in their study.

Appendix C: Interruption Interfaces

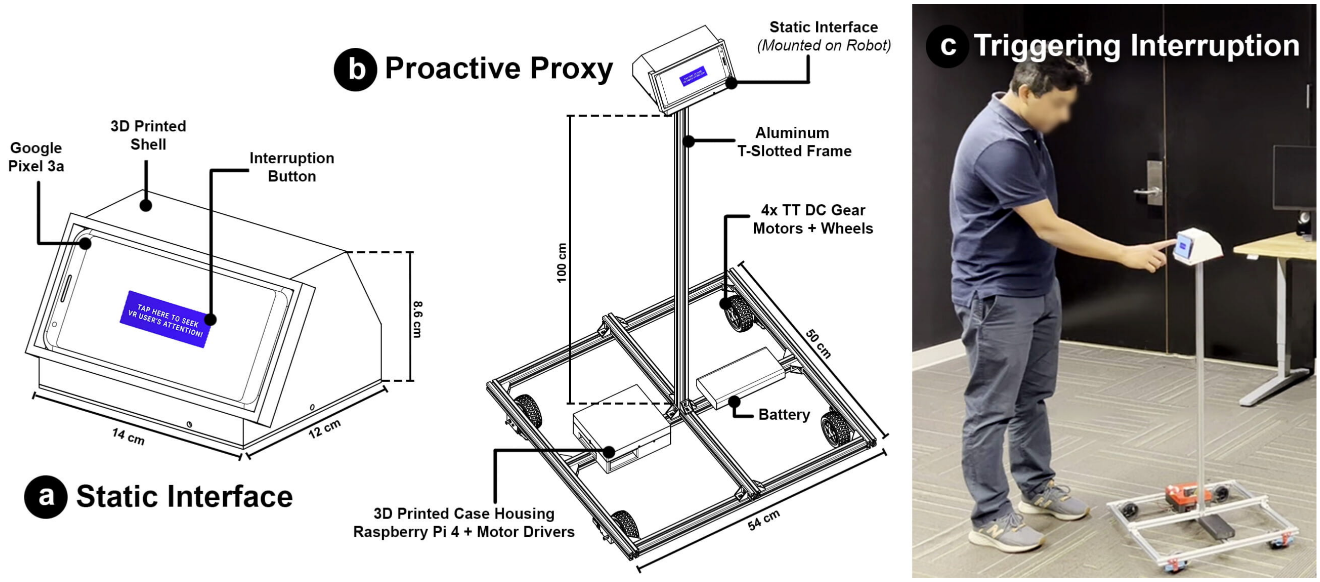

Figure C1: Illustration of the interruption interfaces used during the study. The (a) static interruption interface and (b) the proactive proxy interface. In (c), a user is triggering an interruption using the proactive proxy interface.