O'Hagan, J., Williamson, J. R. , Khamis, M. and McGill, M. (2022) Exploring Manipulating In-VR Audio To Facilitate Verbal Interactions Between VR Users And Bystanders. In: International Conference on Advanced Visual Interfaces (AVI 2022), Rome, Italy, 6-10 June 2022.

This is the author's version of the work. It is posted here for your personal use. You are advised to consult the published version if you wish to cite from it: https://doi.org/10.1145/3531073.3531079

Copyright © 2022 ACM.

http://eprints.gla.ac.uk/269519/

Deposited on: 21 April 2022

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Exploring Manipulating In-VR Audio To Facilitate Verbal Interactions Between VR Users And Bystanders

Joseph O'Hagan
University of Glasgow
United Kingdom
j.ohagan.1@research.gla.ac.uk

Julie R. Williamson
University of Glasgow
United Kingdom
Julie.Williamson@glasgow.ac.uk

Mohamed Khamis
University of Glasgow
United Kingdom
Mohamed.Khamis@glasgow.ac.uk

Mark McGill
University of Glasgow
United Kingdom
Mark.McGill@glasgow.ac.uk

ABSTRACT

Despite recent work investigating how VR users can be made aware of bystanders, few have explored how bystander-VR user interactions may be facilitated by, for example, increasing the user's auditory awareness so they can better converse with bystanders. Through a lab study (N=15) we investigated 4 approaches of manipulating in-VR audio to facilitate verbal interactions between a VR user and bystander: (1) dynamically reducing application volume, (2) removing background audio, (3) removing sound effects and (4) removing all audio. Our results show audio manipulations can be used to significantly improve a VR user's auditory awareness at the cost of reducing sense of presence in VR. They also show most preferred increased awareness be balanced with decreased presence in VR, however, they also identify a subset of participants who prioritised increasing awareness no matter the cost to presence.

CCS CONCEPTS

• **Human-centered computing → Virtual reality**.

KEYWORDS

Virtual Reality, Mixed Reality, Interruptions, Audio, Awareness

ACM Reference Format:
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 (AVI 2022), June 6–10, 2022, Frascati, Rome, Italy.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3531073.3531079

1 INTRODUCTION

Virtual reality (VR) is often used in shared, social settings but interactions between VR users and bystanders (nearby persons who cannot directly interact with the VR user's virtual environment) remain problematic [32, 33]. Central to this is the occlusive nature of VR headsets which introduce significant barriers to a VR user's awareness of and interaction with bystanders [24]. To overcome this, research has begun to investigate technology-mediated awareness systems to increase a VR user's awareness of bystanders and their immediate surroundings [24, 26, 31, 47], variations of which have begun to be included within consumer VR headsets [28, 29].

Yet, at present, the predominant approach of increasing awareness has focused solely on notifying a VR user of bystander existence, that is, to inform the VR user someone is there. Little work has explored beyond this at how a bystander-VR user interactions might be facilitated. For example, prior work has shown most bystander-VR user interactions involve verbal communication in some capacity [33], however, how a VR user's auditory awareness of their surrounding reality can be increased has been largely ignored. Furthermore, recent work has shown that a VR user's application audio (in-VR audio) can be manipulated (e.g. by removing audio such as background music or sound effects) without influencing the user's presence in VR [35, 37, 38]. Therefore, we designed a user study (N=15) to explore how in-VR audio might be manipulated to facilitate a verbal bystander-VR user interactions.

We developed 4 approaches of manipulating in-VR audio to increase a VR user's auditory awareness of a speaking bystander: (1) dynamically lowering volume and removing (2) background audio, (3) sound effect audio and (4) all audio. We evaluated our approaches using a task where participants conversed with the experimenter while playing a target throwing game. We evaluated how each approach facilitated the VR user's ability to converse and impacted sense of presence in VR. Our results found 3 of our audio manipulations (dynamically lowering volume and removing background/all audio) improved awareness and facilitated the interaction. However, this increased awareness was found to come at the cost of significantly lowering the user's sense of presence in VR. While removing sound effects did not significantly disrupt the user's sense of presence it also did little to increase awareness. Our results also show the majority of participants preferred audio manipulations which attempted to balance increasing awareness with maintaining presence with most favouring dynamically lowering the volume. However, we also found a subset of participants who prioritised full awareness of reality (remove all audio) no matter the cost to their presence. We close by discussing our results and pertinent challenges for future work to consider.

AVI 2022, June 6–10, 2022, Frascati, Rome, Italy
© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of the 2022 International Conference on Advanced Visual Interfaces (AVI 2022), June 6–10, 2022, Frascati, Rome, Italy*, https://doi.org/10.1145/3531073.3531079.

2 RELATED WORK

2.1 VR User Awareness of Bystanders

The often problematic nature of bystander-VR user interactions [24, 33] has seen the development of cross-reality awareness systems to increase a VR user's awareness of bystanders. McGill et al. were the first to investigate how a VR user might be automatically informed of bystander existence using contextually displayed photoreal avatars within the VR scene [24]. However, while this approach was found to increase awareness it also significantly disrupted a user's sense of presence in VR and so they concluded alternative, less disruptive, solutions were needed. Building on McGill et al's work, a range solutions have been explored. While work has proposed haptic [13] (e.g. controller vibrations) or audio notifications (e.g. audio alerts) [13, 26, 31], the vast majority have focused on visual solutions such as text notifications [13, 31, 40] or avatars [13, 14, 21, 26, 46]. And while this work has proven a VR user can successfully be made aware when a bystander is co-present, it is limited in that it has focused exclusively on the problem of informing the VR user when someone is there but has not considered how interactions with the bystander might be facilitated beyond this.

For example, prior work has shown most bystander-VR user interactions include some verbal interaction [33], however, little work has explored how a VR user's auditory awareness (their ability to hear their surroundings) might be increased. What work has investigated this is limited to exploring attitudes towards dynamically adjusting application audio to direct the VR user's attention towards nearby sounds of interest [31], although, how effective this approach is at facilitating verbal interactions is unknown. Instead, it is often assumed the on-board, *"acoustically transparent"* audio solutions [25] of existing consumer headsets (e.g. speakers built into the straps of the headset [31]), designed to provide the user with partial auditory awareness of their surroundings, are sufficient at providing increased auditory awareness if desired. However, prior work has highlighted consumer sentiment towards these on-board systems is mixed and the awareness they provide insufficient for some who are forced to adopt alternative solutions such as muting [32]. Additionally, work which investigated bystander-VR user interactions in-the-wild [8, 33] reports empirical evidence of (1) bystander verbal interactions being missed by VR users and (2) VR users removing all audio (either by removing headphones/the headset or muting the application audio) to verbally interact with a bystander. Therefore, how auditory awareness can be increased to facilitate interactions between VR users and bystanders is a pertinent topic of research, yet one that is, at present, unexplored.

2.2 Aural Presence in VR

Audio can serve many functions within an interactive experience [18]. In games audio can provide specific information to a player [20], improve performance [19], affect player behaviour (e.g. increase risk taking [36]) and facilitate immersion/enjoyment [11, 36]. Audio that is perceived as being unsuitable for an experience can also decrease immersion and detract from the user's enjoyment, focus and performance [5, 27, 41]. Therefore, the effects of audio are particularly relevant for VR, where a goal of the technology is to create as immersive an experience as possible for users. Surprisingly, however, recent empirical work has shown minimal effect of background music on user experience in VR in terms of presence [38]. While older literature (late 1990s / early 2000s) strongly suggested ambient noises, sound effects and music all increased immersion/presence in VR [4, 9] studies of *modern VR* (2013 onward) have failed to replicate this [35, 37, 38]. Instead, recent work has indicated audio may be less prominent in creating immersive VR experiences than previously thought and recent work has shown ambient noises, sound effects and background music can be removed from a VR scene without altering the user's presence [35, 37, 38]. While audio designed as a focal point of an experience can still play a significant role in immersion/presence [27, 41], the findings of recent work indicate audio, not designed as a focal point, can be manipulated without altering the user's presence. Therefore, to explore this phenomenon of manipulating in-VR audio without impacting presence and investigate how it might be used to increase a VR user's auditory awareness we designed a user study to evaluate (1) the effectiveness of 4 in-VR audio manipulations at facilitating a verbal bystander-VR user interaction and (2) what impact, if any, they had on the VR user's sense of presence.

3 STUDY OVERVIEW

To explore how audio manipulation could be used to facilitate verbal bystander-VR user interactions, we designed a study (Section 5) where a VR user, playing a game (Section 3.1), conversed with a bystander (an often occurring scenario in-the-wild [8, 33]).

3.1 Our Game's Design

We developed an arcade-style game which participants tasked with throwing cubes at moving targets to score points (similar to [26]). Our game was designed to be a fixed, room-scale experience where users predominantly looked straight ahead and occasionally to their right. This ensured, by design, participants faced the experimenter who sat opposite them 4 metres away, and is an approach used in prior work [24, 26, 46]. As prior work has also shown participant's task can influence attitudes toward bystander awareness systems (e.g. a video watching task vs an active game task [31]) we opted to develop an active experience - one requiring player movement, direct interactions with the virtual environment and was both visually and aurally demanding. Such an experience is more representative of typical, current, consumer VR applications [15, 30] and is more ecologically valid given the reliance on direct user interactions in existing VR games [44].

Our game's task was to throw cubes at moving targets to score points within a fixed time period. This was chosen as a simple, yet effective, way of creating engaging gameplay [45]. 1 point was gained for every target destroyed and 2 points lost if a target self-destructed (was not destroyed quickly enough after spawning). The targets' spawn area was fixed to focus the user's attention forwards (in the direction of the experimenter). To add challenge to the gameplay, a target's design varied by shape (either a cube or cuboid) and movement (either stationary, moving left-to-right in front of the participant or away-and-toward the participant). These parameters were randomly selected from a range decided by the researchers during playtesting of the application. A video demonstration of the game and conditions is shown here[1].

[1]https://youtu.be/YtxuimMgPC4

Our Game's Audio Design: The game audio consisted of two elements: background and sound effect audio. The background audio was persistent, non-diegetic audio present while the game was being played - an upbeat instrumental track. The sound effect audio was a one-off, diegetic sound effect emitted when a target was destroyed - a *"pew"* sound if destroyed by the participant and an electric static *"pist"* sound if it self-destructed.

During development we conducted a small pilot test where 5 individuals were shown the game and asked (1) if the audio matched the game's aesthetic and (2) if they could clearly hear the sound effects over the background audio (e.g. could they determine when a target was destroyed without looking). This was to ensure the audio fit thematically with the game's aesthetic (as mismatches can influence player experience [34]) and the sound effects were noticeable (as one condition involved their removal). All individuals agreed the audio fit thematically and their feedback was used to adjust the audio mix to achieve a suitable balance of background and sound effect audio.

4 DESIGN OF OUR AUDIO MANIPULATIONS

We implemented 4 in-VR audio manipulations derived from findings of prior work. Our approaches were:

- **Dynamic Audio:** Automatically lower VR application audio (to 25% the starting volume) based prior work [31]
- **Remove Sound Effect Audio:** Remove sound effect audio, background audio remains at full volume [18, 35, 37, 38]
- **Remove Background Audio:** Remove background audio, sound effect audio remains at full volume [18, 35, 37, 38]
- **Remove All Audio:** Remove background and sound effect audio (all audio) [33, 39]

Dynamic Audio was included as prior work [31] found it could effectively direct a VR user's attention toward nearby sounds but did not investigate how effectively it could facilitate a bystander-VR user interaction. Removing partial audio (*Remove Sound Effect Audio & Remove Background Audio*) were included to explore how the phenomena of removing audio without altering a VR user's presence [35, 37, 38] might be used to facilitate verbal bystander-VR user interactions. Finally, *Removing All Audio* was included to explore the extreme of fully removing aural presence (the aural equivalent of removing the headset) as prior works have reported muting audio and headphone/headset removal as a behaviour of some VR users during bystander-VR user interactions [32, 33, 39].

We based our implementation on systems built in prior work [31] where the audio manipulation was applied upon detecting external speech. However, as in prior work [14, 26, 31], we used a wizard of oz approach [16] where the manipulation operated on a timer and the verbal interaction was timed to start 1 second before the manipulation was triggered. As in prior work [31], all approaches used a fading effect to decrease/remove the audio over a period of 0.5 seconds and the same effect to increase/return the audio to its original level when exposure to manipulated audio ended.

5 STUDY DESIGN

Our experiment had 5 conditions: 1 for each of the 4 audio manipulations and a baseline condition (no manipulation). Our experiment contained 2 parts: a training and an experimental phase.

The training phase was used to introduce each condition to the participant to ensure they were familiar with all of our approaches before they evaluated any one of them. In this phase, 1 condition (1 session of the game) lasted 50 seconds with 30 seconds exposure to the audio manipulation (starting after 15 seconds). Before starting each condition, the experimenter introduced the condition to the participant (e.g. *"This is the remove all audio condition, all of the audio will be removed, ok start when you are ready"*).

The experimental phase was used to evaluate each condition. In this phase, 1 condition lasted 95 seconds with 60 seconds exposure to the audio manipulation (starting after 25 seconds). Here, when the condition started, the experimenter began a timer and after 24 seconds, just before the audio manipulation triggered, initiated a verbal interaction with the bystander using prepared conversation starters. After the condition ended the participant removed the headset and completed a questionnaire.

5.1 Conversation Starters

In the experimental phase, the experimenter used prepared conversation starters to initiate a verbal interaction with the participant. These were structured so the experimenter made a statement, asked a question then asked follow up questions based on the participant's responses. The experimenter ended the conversation as exposure to audio manipulation ended. The 5 conversation starters were:

- *"I'm thinking of having pizza for dinner later, do you know what you are going to eat for dinner today?"*
- *"I've been listening to a lot of The Beatles lately, have you been listening to anything in particular lately?"*
- *"I've always wanted to go to Egypt to see the pyramids, is there anywhere in the world that you want to see?"*
- *"My full time job is to conduct research into human-computer interaction, what do you do?"*
- *"My favourite colour is purple, what's your favourite colour and why?"*

The experimenter used a decibel meter to monitor the volume of their speaking voice during the experimental phase and first spoke at approximately 60 dB (the average speaking volume [1]). If the participant failed to hear this, the experimenter repeated themselves at approximately 75 dB. If the participant again failed to hear, the experimenter tried again at 75 dB before giving up. The experimenter noted when a participant failed to hear them and any relevant comments made (e.g. *"I can't really hear you"*). Prior to conducting the experiment, the experimenter rehearsed their timing and speaking volume 15 times. Conversation starters were counterbalanced across all participants.

5.2 Questionnaire Measures

We designed a questionnaire to evaluate (1) how effective our audio manipulations were at facilitating a verbal bystander-VR user interaction, (2) their impact on sense of presence and (3) their usability. All questions used a 7-point Likert scale. We did not ask (3) for the baseline condition as the questions were not applicable.

(1) Facilitating Verbal Interactions Questions: we evaluated participants experience during the verbal interaction using 6 questions - to what extent they agreed (1=strongly disagree, 7=strongly agree): (1) *"It felt as if you and the person you heard were together* *in the same place"*, (2) *"I could successfully converse with the experimenter"*, (3) *"I could clearly hear the experimenter"*, (4) *"Talking, whilst wearing the headset, felt uncomfortable"*, (5) *"The experimenter was easy to understand"* and (6) *"The conversation felt natural"*.

(2) Presence Questions: we evaluated presence using the *"Sense of Being There"* subset of the IPQ questionnaire [42] and 2 questions derived from similar questions asked in prior works [13, 24, 26, 31, 43, 46] - to what extent participants agreed (1=strongly disagree, 7=strongly agree) (1) *"I enjoyed my experience in VR"* and (2) *"I was too aware of my real world surroundings"*.

(3) Usability Questions: we evaluated usability using 7 questions - to what extent participants agreed (1=strongly disagree, 7=strongly agree) the audio manipulation: (1) *"was disruptive"*, (2) *"was frustrating"*, (3) *"was urgent"*, (4) *"felt natural"*, (5) *"was easy to understand"*, (6) *"was informative"* and (7) *"improved their ability to verbally interact with a bystander"*.

5.3 Limitations

We investigated the context of a VR user playing a game and verbally interacting with a known bystander in a private setting. However, as context can influence attitudes towards how/when/why awareness is increased [31] further work is needed to explore alternative contexts like a VR user using a productivity application, an interaction within a public space [2, 22, 47], etc. We also opted to use the on-board audio system of the Quest 2 headset to investigate its effectiveness at facilitating verbal interactions. Over/in-ear headphones are, however, often used by VR users [32] and likely have some influence, especially if designed to block out a user's surrounding reality (e.g. noise cancelling). Finally, we assumed the onset of the audio manipulation would be detection of external speech. However, alternative triggers could be used (e.g. detecting someone is there) which also may influence participant responses.

5.4 Procedure

Upon arrival the experiment's purpose was explained and a consent form and demographic questionnaire given to the participant. Participants were told they would be playing a VR game and would experience 4 auditory awareness systems we had designed to improve their ability to verbally interact with a non-VR bystander. It was explained the experiment would consist of 2 parts: a first part to introduce each awareness system and a second to evaluate each approach. Participants were told during the second part the experimenter would act as a known bystander and verbally interact with them, that they were free to respond as they wished to this, and it was fine if they did not hear the experimenter. A demonstration video of the game was then shown and its controls explained. Participants were then instructed where to stand during the experiment and shown (if required) how to put on and fit the headset. The experimenter took their position seated 4 metres away facing the participant on the opposite side of the room.

Participants then began the training phase. During this phase, participants were told to set the system volume of the headset to a comfortable but immersive level. Most set it to about 60% system volume - a level which meant the experimenter could also clearly hear it being emit from the headset. After the training phase was completed participants were instructed to take off the headset and were given the opportunity for a break while the experimenter set up the experimental phase. Once ready, the participant and experimenter resumed their positions and began the experimental phase where each condition was evaluated in turn. After all conditions were evaluated, participants were asked to rank order the 4 awareness approaches from best to worse and to describe how they ranked their preferences (e.g. prioritising immersion, awareness, etc). Finally, participants were asked if they had any comments regarding any of the conditions they had experienced.

The experiment took on average 25 minutes to complete. Participants were paid a £5 Amazon voucher for participating. Condition order was counterbalanced across all participants. An Oculus Quest 2 headset was used to conduct the experiment.

6 ANALYSIS

We analysed our data by first calculating the mean and standard deviation values for each of our Likert scale questions. We then used a Friedman test to find significant differences between factors and performed pairwise comparisons using Wilcoxon Signed Rank tests with Bonferroni corrected p-values. For participants' preference ranking, the average ranking score was calculated for each condition. Participants' comments, used to justify their ranking, were coding using initial coding [6] where participants' statements were assigned emergent codes over repeated cycles with the codes grouped using a thematic approach. A single coder performed the coding and reviewed/discussed the coding with one other researcher. Analysis of the observations recorded by the experimenter followed a similar approach where the experimenter's notes were coded then reviewed/discussed with another researcher.

6.1 Participant Demographic Data

Participants were recruited using social media and mailing lists. 15 participants completed our study (7 female, 8 male) aged between 19 and 40 years of age (M=24.8 , SD=5.48). Participants were asked to indicate prior experience with VR using a 5-point Likert scale (1=a little, 5=a lot), (M=3.73, SD=1.22). As our task involved a verbal component we asked if participants were a native English speaker (10 yes, 5 no) and to rate their English proficiency on a 5-point Likert scale (1=basic, 5=native), (M=4.67, SD=0.49).

7 RESULTS

7.1 Observations During Verbal Interaction

Observations made by the experimenter are summarised in Table 2. *Baseline* and *Remove Sound Effect* were the most problematic conditions. Both had 1 participant fail to hear all attempted interactions and multiple participants requiring the experimenter speak louder to be heard. This is somewhat expected, as these conditions provided the lowest amount of awareness, and is in-line with findings of prior work which reported empirical evidence of failed verbal interactions with VR users using on-board audio systems [8, 33]. The similarity experiencing the *Baseline* and *Remove Sound Effect* conditions was also commented on by 5 participants who said (regarding the *Remove Sound Effect* condition) they *P10: "didn't notice what it did to the audio at all"*. Issues, however, were not exclusive to these conditions as 2 participants also required the experimenter speak louder when using the *Dynamic Audio* condition.

Table 1: Mean (standard deviation) values for the facilitating interaction questions (7-point Likert scale: 1=strongly disagree, 7=strongly agree). *Remove Background* and *Remove All Audio* increased awareness the most, closely followed by *Dynamic Audio.*

Facilitating Interaction Questions Mean (Std) Values	*Baseline*	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
(1) Felt together in same place	3.6 (1.58)	5.4 (1.2)	5.93 (0.68)	3.93 (1.91)	5.8 (0.91)
(2) Could successfully converse	2.13 (1.02)	5.6 (1.5)	6.33 (0.7)	3.33 (1.62)	6.27 (1.12)
(3) Could clearly hear the experimenter	1.67 (0.7)	5.6 (1.62)	6.27 (0.85)	2.8 (1.64)	6.67 (0.6)
(4) Talking, wear the headset, was uncomfortable	4.47 (1.59)	3.47 (1.82)	2.4 (1.25)	3.73 (1.91)	2.8 (1.68)
(5) Experimenter was easy to understand	2.00 (0.73)	5.53 (1.59)	6.27 (0.85)	3.33 (1.7)	6.53 (0.72)
(6) Conversation felt natural	3.4 (1.5)	5.27 (1.29)	5.87 (0.88)	3.67 (1.81)	5.4 (1.25)

Table 2: Observations made during the experimental phase

Observation	***Count***	***Percentage of Participants***
— Baseline (No Manipulation)		
Missed attempted interaction	1	6.67%
Repeated self louder to be heard	5	33.33%
— Remove Sound Effects		
Missed attempted interaction	1	6.67%
Repeated self louder to be heard	4	26.67%
— Dynamic Audio		
Repeated self louder to be heard	2	13.33%

Table 3: Significant differences for facilitating the interaction questions. *Remove All Audio, Remove Background* and *Dynamic Audio* all improved the user's awareness

(1) Felt together in the same place	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	**$p<0.005$**	**$p<0.005$**	$p>0.005$	**$p<0.005$**
Dynamic Audio	-	$p>0.005$	**$p<0.005$**	$p>0.005$
Remove Background	-	-	**$p<0.005$**	$p>0.005$
Remove Sound Effects	-	-	-	**$p<0.005$**

(2) Could successfully converse	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	**$p<0.005$**	**$p<0.005$**	$p>0.005$	**$p<0.005$**
Dynamic Audio	-	$p>0.005$	**$p<0.005$**	$p>0.005$
Remove Background	-	-	**$p<0.005$**	$p>0.005$
Remove Sound Effects	-	-	-	**$p<0.005$**

(3) Could clearly hear the experimenter	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	**$p<0.005$**	**$p<0.005$**	$p>0.005$	**$p<0.005$**
Dynamic Audio	-	$p>0.005$	**$p<0.005$**	$p>0.005$
Remove Background	-	-	**$p<0.005$**	$p>0.005$
Remove Sound Effects	-	-	-	**$p<0.005$**

(4) Talking was uncomfortable	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	$p>0.005$	**$p<0.005$**	$p>0.005$	$p>0.005$
Dynamic Audio	-	$p>0.005$	$p>0.005$	$p>0.005$
Remove Background	-	-	$p>0.005$	$p>0.005$
Remove Sound Effects	-	-	-	$p>0.005$

(5) Experimenter was easy to understand	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	**$p<0.005$**	**$p<0.005$**	$p>0.005$	**$p<0.005$**
Dynamic Audio	-	$p>0.005$	**$p<0.005$**	$p>0.005$
Remove Background	-	-	**$p<0.005$**	$p>0.005$
Remove Sound Effects	-	-	-	**$p<0.005$**

(6) The conversation felt natural	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	**$p<0.005$**	**$p<0.005$**	$p>0.005$	**$p<0.005$**
Dynamic Audio	-	$p>0.005$	**$p<0.005$**	$p>0.005$
Remove Background	-	-	**$p<0.005$**	$p>0.005$
Remove Sound Effects	-	-	-	**$p<0.005$**

7.2 Facilitating Verbal Interactions Results

Dynamic Audio, *Remove Background* and *Remove All Audio* all significantly improved participants ability to engage in a verbal interaction. *Remove Sound Effects* had minimal impact and was instead comparable with the *Baseline* (Table 1, Table 3).

Significant differences were found in all questions used during our evaluation (Table 3). The same 6 significant difference pairings were found across questions: *(1) felt together in the same place*, *(2) could successfully converse*, *(3) could clearly hear*, *(5) experimenter was easy to understand* and *(6) conversation felt natural.* These were significant differences between the *Baseline* and the *Dynamic Audio*, *Remove Background* and *Remove All Audio* conditions and significant differences between *Remove Sound Effects* and *Dynamic Audio*, *Remove Background* and *Remove All Audio* conditions. This suggests the *Baseline* and *Remove Sound Effects* conditions are comparable and the removal of sound effects, in our experiment, made no difference to our participants verbal interaction.

This result is also reflected in the mean scores (Table 1) where the *Baseline* performed worst, *Remove Sound Effects* only marginally better, and *Dynamic Audio*, *Remove Background* and *Remove All Audio* all significantly better. While somewhat anticipated, one would expect as more audio is removed the user would be able to hear better, our results do show partial removal (e.g. *Remove Background*) or reduction (e.g. *Dynamic Audio*) of audio can facilitate the interaction with comparable effectiveness as removing all audio.

Finally, regarding the question: *(4) talking, wearing the headset, was uncomfortable* - while the *Baseline* performed worst (m=4.47) its mean score still indicated participants found it *"it neither uncomfortable nor comfortable"*. All other conditions scored lower indicating participants were comfortable with them. Only 1 significant difference was found, however, between the *Baseline* and *Remove Background* conditions, although, the difference between these conditions does highlight the positive benefit bystander awareness systems can have for some VR users' comfort. For the *Baseline*, 53.33% of participants agreed (13.33% weakly agree, 40% agree) they were uncomfortable engaging in the verbal interaction whilst wearing the headset. However, for *Remove Background*, this reduced to only 6.67% agreeing they were uncomfortable which highlights how a noticeable drop in discomfort can be made by the audio manipulation and the awareness it provides.

7.3 Presence Evaluation Results

Dynamic Audio, Remove Background and *Remove All Audio* all significantly decreased sense of presence in VR. *Remove Sound Effects* did not and was comparable to the *Baseline* (Table 4, Table 6).

For *(1) IPQ: Sense of Presence* we found 5 significant differences: 3 between the *Baseline* and *Dynamic Audio*, *Remove Background* and *Remove All* conditions and 2 between *Remove Sound Effects* and

Table 4: Mean (standard deviation) values for the presence questions (7-point Likert scales, 1=strongly disagree, 7=strongly agree) where for the IPQ subset a higher value indicates a greater sense of presence in VR. More substantial audio manipulations (*Remove Background, Remove All Audio*) lowered the presence score but did not alter enjoyment of the experience.

Presence Questions Mean (Std) Values	*Baseline*	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
(1) IPQ: Sense of Presence	5.15 (0.98)	4.11 (0.83)	3.55 (0.91)	4.87 (0.97)	3.6 (1.07)
(2) Enjoyed the VR experience	5.33 (1.58)	5.6 (1.53)	5.33 (1.53)	5.6 (1.45)	5.2 (1.68)
(3) Was too aware of my real world surroundings	1.93 (1.24)	2.4 (1.58)	3.13 (1.71)	1.8 (0.65)	3.4 (1.85)

Table 5: Mean (standard deviation) values for our usability questions. *Dynamic Audio, Remove Background* and *Remove All Audio* performed well, albeit somewhat more disruptive than *Remove Sound Effects*, although this did not cause a rise in frustrating.

Usability Questions Mean (Std) Values	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
(1) Was disruptive	2.67 (1.35)	3.53 (1.5)	1.93 (1.29)	4.27 (1.69)
(2) Was frustrating	2.4 (1.45)	2.2 (1.11)	2.8 (2.01)	3.07 (1.53)
(3) Was urgent	2.13 (0.81)	3.67 (1.62)	1.6 (1.02)	4.67 (1.78)
(4) Felt natural	4.93 (1.48)	5.8 (0.91)	3.2 (2.14)	5.0 (1.15)
(5) Was easy to understand	5.73 (0.93)	5.8 (0.91)	2.33 (1.53)	6.27 (0.93)
(6) Was informative	4.67 (1.45)	4.93 (1.84)	1.8 (1.33)	5.53 (1.2)
(7) Improved ability to verbally interact	6.07 (0.93)	6.6 (0.61)	2.27 (1.48)	6.4 (0.71)

Table 6: Significant differences for the presence questions. *Remove All/Background Audio* and *Dynamic Audio* significantly disrupted sense of presence but not enjoyment.

(1) IPQ: Sense of Presence	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	***p<0.005***	***p<0.005***	*p>0.005*	***p<0.005***
Dynamic Audio	-	*p>0.005*	***p<0.005***	*p>0.005*
Remove Background	-	-	*p>0.005*	*p>0.005*
Remove Sound Effects	-	-	-	***p<0.005***

(2) Enjoyed the VR experience	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	*p>0.005*	*p>0.005*	*p>0.005*	*p>0.005*
Dynamic Audio	-	*p>0.005*	*p>0.005*	*p>0.005*
Remove Background	-	-	*p>0.005*	*p>0.005*
Remove Sound Effects	-	-	-	*p>0.005*

(3) Was too aware of real world	*Dynamic Audio*	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Baseline	*p>0.005*	*p>0.005*	*p>0.005*	*p>0.005*
Dynamic Audio	-	*p>0.005*	*p>0.005*	*p>0.005*
Remove Background	-	-	*p>0.005*	*p>0.005*
Remove Sound Effects	-	-	-	*p>0.005*

the *Dynamic Audio* and *Remove All Audio* conditions (Table **??**). No significant difference was found between the *Baseline* and *Remove Sound Effects* reinforcing their similarity. As is somewhat expected, more substantial audio manipulations (e.g. *Dynamic Audio*, *Remove Background*, *Remove All Audio*) caused a higher decrease in presence, although, *Dynamic Audio* (m=4.11) retained more presence than both *Remove Background* (m=3.55) and *Remove All Audio* (m=3.6).

Despite this decrease in presence, participants enjoyment varied minimally across the conditions (mean scores ranging from 5.2 to 5.6) and no significant difference was found between conditions. Furthermore, participants did not consider any condition to make them *"too aware"* of their surrounding reality. As such, although mean scores increased as more substantial audio manipulations were made no condition was said to increase awareness too much.

7.4 Usability Evaluation Results

Our usability question's mean scores are summarised in Table 5 and the statistical differences between the pairwise comparisons in Table 7. Each usability question is discussed, in turn, next.

7.4.1 ***Disruptive:*** Our more substantial audio manipulations caused higher levels of disruption to participant's experience in VR, although *Remove All Audio* (m=4.27) was the only condition considered disruptive. 3 significant differences were found: between *Remove All Audio* and the *Dynamic Audio* and *Remove Sound Effects* conditions and between *Remove Background* and *Remove Sound Effects*. Noteworthy, is *Dynamic Audio* (m=2.67) which, similar to its performance in sense of presence, was not considered as disruptive/impactful to the experience as the *Remove Background* and *Remove All Audio* conditions.

7.4.2 ***Frustrating:*** No approach was said to be frustrating - a positive result for all of our approaches. No significant differences were found between the conditions either. Similar to disruption, *Remove All Audio* (m=3.07) scored highest although it was still considered not frustrating by our participants.

7.4.3 ***Urgency:*** Our more substantial audio manipulations were considered more urgent than the other approaches. 3 significant differences were found: between *Remove All Audio* and the *Dynamic Audio* and *Remove Sound Effects* conditions and between *Remove Background* and *Remove Sound Effects*. As with disruption, no significant difference was found between *Dynamic Audio* and *Remove Sound Effect* reinforcing *Dynamic Audio* was considered by our participants to be somewhat lesser than *Remove Background* and *Remove All Audio* in its intrusiveness.

7.4.4 ***Natural:*** *Dynamic Audio*, *Remove Background* and *Remove All Audio* were said to be natural ways increasing awareness. *Remove Sound Effects* was not, likely because participants regarded it as insufficient for increasing awareness. *Remove Background* was considered the most natural with comments made by 3 participants providing some insight into why when they described it as *P15: "the most obvious attempt at reducing audio but maintaining presence"*. These participants felt awareness should be increased while maintaining presence in VR and believed *Remove Background* most obviously attempted this. *Dynamic Audio*, in contrast, had the same purpose but perhaps achieved it in a less noticeable manner. 2 significant differences were found between *Remove Sound Effects* and the *Remove Background* and *Remove All* conditions.

7.4.5 *Easy to Understand:*
Dynamic Audio, *Remove Background* and *Remove All Audio* were well understood by our participants. *Remove Sound Effects* was not, likely because it frequently went unnoticed. During the experiment 5 participants commented on this stating they *"didn't notice"* how it differed from the *Baseline*. 3 significant differences were found between *Remove Sound Effect* and all of the other conditions.

7.4.6 *Informative:*
An increase in informativeness was also seen as more substantial audio manipulations were made (which is somewhat expected, e.g. the more audio removed, the more aware of the surrounding reality the user will be). *Dynamic Audio* again performed comparable to *Remove Background* and *Remove All Audio*. 3 significant differences were found between *Remove Sound Effect* and all of the other conditions.

7.4.7 *Improved Ability To Verbally Interact:*
Dynamic Audio, *Remove Background* and *Remove All Audio* were said to improve participants' ability to verbally interact. All performed comparably and the results are similar to those seen our questions investigating our conditions effectiveness at facilitating verbal interactions (Section 7.2). As in those results *Remove Sound Effects* was not considered sufficient for facilitating verbal interactions while the other conditions were. 3 significant differences were found: between *Remove All Audio* and *Dynamic Audio* and *Remove Sound Effects* and between *Remove Background* and *Remove Sound Effects*.

7.5 Preference Ranking Results

The average ranking score of participants preference ranking is shown in Table 8. *Dynamic Audio* scored highest (3.27 out of 4) and was the 1st choice of 60% of participants. *Remove Sound Effects* performed worst (1.13 out of 4) which is expected given its performance across the other evaluation metrics. Interestingly, *Remove Background* and *Remove All Audio* both scored 2.8, although the composition of their scores differs. *Remove Background* was favoured primarily as a 2nd choice whereas *Remove All Audio* was spread more uniformly across the 1st, 2nd and 3rd choices.

When justifying their ranking, 11 participants said they attempted to balance increased awareness with maintaining presence in VR. 9 of the 11 selected *Dynamic Audio* as their first choice as they viewed it *P1: "the best compromise of awareness and immersion"*. The 2 others selected *Remove Background* as first choice as they were *P12: "slightly in favour of prioritising awareness"*. The remaining 4 participants all ranked *Remove All Audio* as their first choice and indicated their only concern was increasing their awareness regardless of the cost to presence, *P10: "I want awareness and don't care what my immersion is like at that point"*.

8 DISCUSSION & FUTURE WORK

8.1 Supporting The Varying Wants of Users

The majority of our participants (11 of 15) indicated they preferred increased awareness be balanced with retention of presence in VR. For most *Dynamic Audio* best fit this aim. The results of our evaluation reinforce this as *Dynamic Audio* was found to perform comparable to *Remove Background* and *Remove All Audio* at facilitating the interaction while being slightly less disruptive to the user's experience. However, a subset of our participants (4 of 15) said they were uninterested in balancing awareness/presence, rather, they wanted to fully prioritise increasing awareness no matter the cost to presence (e.g. *Remove All Audio*). This differing attitude toward why/how to increase awareness is not unexpected, however, as similar sentiments have been reported throughout prior works [12, 14, 21, 31, 46] with Medeiros et al concluding in their own bystander awareness study *"there exists no one-size-fits-all approach"* to increasing VR user awareness [26].

Consequentially, it is likely a range of awareness systems, providing varying levels of awareness/presence, need be available to the VR user. At a minimum our results indicate a need to offer *no awareness* (e.g. no manipulation), *full awareness* (e.g. *Remove*

Table 7: Significant differences for our usability questions. Most differences were found with *Remove Sound Effects*.

(1) Was disruptive	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	*p>0.0083*	***p<0.0083***
Remove Background	-	***p<0.0083***	*p>0.0083*
Remove Sound Effects	-	-	***p<0.0083***

(2) Was frustrating	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	*p>0.0083*	*p>0.0083*
Remove Background	-	*p>0.0083*	*p>0.0083*
Remove Sound Effects	-	-	*p>0.0083*

(3) Was urgent	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	*p>0.0083*	***p<0.0083***
Remove Background	-	***p<0.0083***	*p>0.0083*
Remove Sound Effects	-	-	***p<0.0083***

(4) Felt natural	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	*p>0.0083*	*p>0.0083*
Remove Background	-	***p<0.0083***	*p>0.0083*
Remove Sound Effects	-	-	***p<0.0083***

(5) Was easy to understand	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	***p<0.0083***	*p>0.0083*
Remove Background	-	***p<0.0083***	*p>0.0083*
Remove Sound Effects	-	-	***p<0.0083***

(6) Was informative	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	***p<0.0083***	*p>0.0083*
Remove Background	-	***p<0.0083***	*p>0.0083*
Remove Sound Effects	-	-	***p<0.0083***

(7) Improved ability to verbally interact	*Remove Background*	*Remove Sound Effects*	*Remove All Audio*
Dynamic Audio	*p>0.0083*	***p<0.0083***	*p>0.0083*
Remove Background	-	***p<0.0083***	*p>0.0083*
Remove Sound Effects	-	-	***p<0.0083***

Table 8: Preference of conditions (average ranking score out of a maximum of 4). *Dynamic Audio* was the preferred approach, *Remove Background* and *Remove All Audio* were tied second although the composition of their scores varies.

Condition / Preference	***1st***	***2nd***	***3rd***	***4th***	***Average Ranking***
Dynamic Audio	9	2	3	1	3.27
Remove Background	2	8	5	0	2.8
Remove Sound Effects	0	0	2	13	1.13
Remove All Audio	4	5	5	1	2.8

All Audio) and *balanced awareness* (e.g. *Dynamic Audio*) options. Ideally, this *balanced awareness* option would allow VR users to specify the awareness/presence balance provided. For example, our *Dynamic Audio* approach automatically lowered the in-VR audio to 25% of the starting volume but this could be easily configured by the user to a lower or higher value (e.g. 10% or 75%) via system settings to provide a different awareness/presence balance. Similarly, our approaches which partially removed audio (*Remove Sound Effects* and *Remove Background*) could be options within a larger system to allow users to specify which audio elements within an application to reduce/remove when providing increased awareness. Applications, such as video games, already allow users to adjust their audio balance (e.g. altering the volume of background music, sound effects, dialogue, etc) as an accessibility feature [38] and one can envision how similar systems could be used to allow users to specify how awareness is increased. These, however, represent only 2 ways in which a balanced awareness/presence could be provided and further work is needed to investigate/explore alternative approaches.

8.2 Presence & Preference

We found audio manipulations which both did and did not significantly alter the user's sense of presence. However, our approach which did not significantly impact sense of presence (*Remove Sound Effects*) proved ineffective at facilitating verbal interactions. *Dynamic Audio* does offer some consolation, however, as it was participants preferred approach, was effective at facilitating the interaction and retained more presence than *Remove Background*/*Remove All Audio*. Furthermore, like all our approaches, it did not decrease participant's enjoyment of their experience in VR.

It is worth highlighting then that although *"impact on sense of presence"* is a beneficial measure to capture when evaluating bystander awareness systems it represents only one factor and should not be fixated on above others. This is a limitation of some prior works which assume awareness preference correlates with minimal disruption to presence, meaning, the preferred awareness approach is the one which provides awareness whilst retaining as much presence as possible [13, 24, 46]. While true for some this does not represent all users. For example, 4 of our participants indicated wanted to fully prioritise awareness no matter the impact to their presence. Prior work has reported similar behaviour when exploring notifying a VR user of bystander existence where some VR users were made uncomfortable if told someone was nearby but not where and so preferred a more disruptive notification because of additional information provided [31]. For these individuals, retained presence was not justified by the loss in awareness as the reduced awareness created increased anxiety/frustration which then impacted enjoyment of the experience [31]. Therefore, while sense of presence is an important measure to capture, it is only one of many possible measures and should not be fixated on above others when evaluating a bystander awareness system.

8.3 Pertinent Challenges For Future Work

While much progress has been made towards building systems to automatically increase a VR user's awareness of their surrounding reality there remains many challenges for future work to address. In our work, we focused on a single context - a known bystander, private setting, VR user playing a game - and so further work is necessary to explore alternative contexts (e.g. a VR user using a productivity application, interactions in the workplace or public spaces, etc). Furthermore, our study, like prior works, used a lab study but future work is necessary to explore awareness systems in-the-wild (e.g. via diary or remote studies [23]) to obtain empirical evidence of usage on much larger sample sizes. We also focused solely on manipulating in-VR audio but alternative approaches (used alongside or independent from audio manipulation) such as amplifying bystander speech [10] or generating real-time subtitles [7] may also prove effective. These more intrusive methods require exploration of user attitudes/comfort towards them also [3].

Furthermore, the role of increasing auditory awareness has yet to be fully established within the corpus of bystander awareness systems work. Awareness can be increased in many ways [17] and consumer VR headsets will be capable of increasing the user's awareness in a variety of them [13, 31]. Some systems will focus on providing visual awareness of who is there, where they are, etc, while others will focus on increasing auditory awareness of nearby sounds, facilitating verbal interactions, etc. Yet, at present, it is unclear what it means to increase awareness using one approach over another. Existing work cannot say, for example, given a range bystander awareness systems (each known to increase awareness differently) which a VR user would use, when and why. While works [12, 24] have proposed dynamically adjusting awareness with the VR user's engagement with the bystander, no work has investigated this using a range of awareness systems across a range of bystander-VR user interaction scenarios. Exploring this, and the use of bystander awareness systems more broadly, is essential going forward. While prior works have proven a VR user can be automatically notified when someone is nearby, and we in this paper have shown how verbal interactions can be facilitated, it remains unknown how these systems will be used by users in practice.

9 CONCLUSION

Despite prior work investigating how to automatically inform a VR user of bystanders, few have explored how to improve bystander-VR user interactions beyond notification. Through a lab study we investigated 4 approaches of manipulating in-VR audio to facilitate a verbal interaction between a VR user and bystander. Our results show audio manipulations can significantly improve a VR user's awareness and enhance their experience during verbal interactions. They also show most preferred increased awareness be balanced with decreased presence, believing dynamically lowering the application volume was our best approach at achieving this. However, our results also identify a subset of participants whose preference was to prioritise increased awareness no matter the cost to presence, highlighting no one-size-fits-all approach can satisfy all users.

REFERENCES

[1] Decibel Pro App. 2021. How Many Decibels Does A Human Speak Normally. https://decibelpro.app/blog/how-many-decibels-does-a-human-speak-normally/. Accessed: 2022-01-04.
[2] L. Bajorunaite, S. Brewster, and J. R. Williamson. 2021. Virtual Reality in transit: how acceptable is VR use on public transport?. In *IEEE Conf. on VR and 3D User Interfaces*. 432–433. https://doi.org/10.1109/VRW52623.2021.00098

[3] Jolie Bonner, Joseph O'Hagan, Florian Mathis, Jamie Ferguson, and Mohamed Khamis. 2021. Using Personal Data to Support Authentication: User Attitudes and Suitability. In *20th Conference on Mobile and Ubiquitous Multimedia (MUM 2021)*. ACM, New York, NY, USA, 35–42. https://doi.org/10.1145/3490632.3490644

[4] Sarah Brown, Ilda Ladeira, Cara Winterbottom, and Edwin Blake. 2003. The effects of mediation in a storytelling virtual environment. In *International Conference on Virtual Storytelling*. Springer, 102–111.

[5] G. Cassidy and R. MacDonald. 2009. The effects of music choice on task performance: A study of the impact of self-selected and experimenter-selected music on driving game performance and experience. *Musicae Scientiae* (2009).

[6] Strauss A. L. Corbin J. M. 1998. *Basics of qualitative research: techniques and procedures for developing grounded theory*. SAGE Publications, Inc.

[7] Itai Dabran, Tzoof Avny, Eytan Singher, and Haim Ben Danan. 2017. Augmented reality speech recognition for the hearing impaired. In *2017 IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems*.

[8] Emily Dao, Andreea Muresan, Kasper Hornbæk, and Jarrod Knibbe. 2021. Bad Breakdowns, Useful Seams, and Face Slapping: Analysis of VR Fails on YouTube. In *Proc. of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, USA. https://doi.org/10.1145/3411764.3445435

[9] H. Q. Dinh, N. Walker, L. F. Hodges, C. Song, and A. Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE VR (Cat. 99CB36316)*. IEEE, 222–228.

[10] Lionel Fontan, Maxime Le Coz, Charlotte Azzopardi, Michael A Stone, and Christian Füllgrabe. 2020. Improving hearing-aid gains based on automatic speech recognition. *The Journal of the Acoustical Society of America* 148, 3 (2020).

[11] Jiulin Zhang Xiaoqing Fu. 2015. The influence of background music of video games on immersion. *Journal of Psychology & Psychotherapy* 5, 4 (2015).

[12] C. George, A. N. Tien, and H. 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)*. 412–424. https://doi.org/10.1109/ISMAR50242.2020.00067

[13] S. Ghosh, L. Winston, N. Panchal, P. Kimura-Thollander, J. Hotnog, D. Cheong, G. Reyes, and G. D. Abowd. 2018. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1447–1456. https://doi.org/10.1109/TVCG.2018.2793698

[14] Matthew Gottsacker, Nahal Norouzi, Kangsoo Kim, G. Bruder, and Greg Welch. 2021. Diegetic Representations for Seamless Cross-Reality Interruptions.

[15] M.P. Jacob Habgood, David Wilson, David Moore, and Sergio Alapont. 2017. HCI Lessons From PlayStation VR. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 125–135. https://doi.org/10.1145/3130859.3131437

[16] Bruce Hanington and Bella Martin. 2019. *Universal methods of design expanded and revised: 125 Ways to research complex problems, develop innovative ideas, and design effective solutions*. Rockport publishers.

[17] Scott E. Hudson and Ian Smith. 1996. Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems. In *Proc. of the 1996 ACM Conference on Computer Supported Cooperative Work (CSCW '96)*. ACM, New York, NY, USA. https://doi.org/10.1145/240080.240295

[18] D. Jain, S. Junuzovic, E. Ofek, M. Sinclair, J. Porter, C. Yoon, S. Machanavajhala, and Meredith R. M. 2021. *A Taxonomy of Sounds in Virtual Reality*. ACM, New York, NY, USA, 160–170. https://doi.org/10.1145/3461778.3462106

[19] Colby Johanson and Regan L Mandryk. 2016. Scaffolding player location awareness through audio cues in first-person shooters. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3450–3461.

[20] Kristine Jørgensen. 2008. *Left in the dark: playing computer games with the sound turned off*. Ashgate.

[21] Yoshiki Kudo, Anthony Tang, Kazuyuki Fujita, Isamu Endo, Kazuki Takashima, and Yoshifumi Kitamura. 2021. Towards Balancing VR Immersion and Bystander Awareness. *Proc. ACM Hum.-Comput. Interact.* 5, ISS, Article 484 (nov 2021), 22 pages. https://doi.org/10.1145/3486950

[22] Jingyi Li, Filippe Frulli, Stella Clarke, and Andreas Butz. 2022. Towards Balancing Real-World Awareness and VR Immersion in Mobile VR. (2022).

[23] Florian Mathis, Xuesong Zhang, Joseph O'Hagan, Daniel Medeiros, Pejman Saeghe, Mark McGill, Stephen Brewster, and Mohamed Khamis. 2021. Remote XR Studies: The Golden Future of HCI Research?. In *CHI 2021 Workshop on XR Remote Research*. http://www.mat.qmul.ac.uk/xr-chi-2021/

[24] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proc of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA. https://doi.org/10.1145/2702123.2702382

[25] Mark McGill, Stephen Brewster, David McGookin, and Graham Wilson. 2020. Acoustic transparency and the changing soundscape of auditory mixed reality. In *Proc. of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–16.

[26] D. Medeiros, R. Anjos, N. Pantidi, K. Huang‡, M. Sousa, C. Anslow, and J. Jorge. 2021. Promoting Reality Awareness in Virtual Reality through Proxemics. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 21–30.

[27] Adrian C North and David J Hargreaves. 2000. Musical preferences during and after relaxation and exercise. *American journal of psychology* 113, 1 (2000), 43–68.

[28] Oculus. 2019. Oculus Guardian. https://developer.oculus.com/documentation/native/pc/dg-guardian-system/. Accessed: 2020-09-01.

[29] Oculus. 2021. Space Sense Update. https://www.oculus.com/blog/a-smarter-guardian-android-phone-notifications-and-improved-voice-commands-in-latest-oculus-software-update. Accessed: 2021-11-03.

[30] Joseph O'Hagan, Mohamed Khamis, and Julie R. Williamson. 2021. Surveying Consumer Understanding & Sentiment Of VR. In *Proc of the International Workshop on Immersive Mixed and Virtual Environment Systems* (Istanbul, Turkey). ACM, New York, NY, USA, 14–20. https://doi.org/10.1145/3458307.3460965

[31] Joseph O'Hagan and Julie R. Williamson. 2020. Reality Aware VR Headsets. In *Proc. of the 9TH ACM International Symp. on Pervasive Displays (PerDis '20)*. ACM, New York, NY, USA, 9–17. https://doi.org/10.1145/3393712.3395334

[32] Joseph O'Hagan, Julie R. Williamson, and Mohamed Khamis. 2020. Bystander Interruption of VR Users. In *Proc. of the 9TH ACM International Symp. on Pervasive Displays*. ACM, New York, NY, USA. https://doi.org/10.1145/3393712.3395339

[33] 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 Symp. on Mixed and Augmented Reality (ISMAR)*. IEEE Computer Society.

[34] Giovanni Ribeiro, Katja Rogers, Maximilian Altmeyer, Thomas Terkildsen, and Lennart E Nacke. 2020. Game Atmosphere: Effects of Audiovisual Thematic Cohesion on Player Experience and Psychophysiology. In *Proc of the Annual Symp. on Computer-Human Interaction in Play*.

[35] K. Rogers, M. Jörg, and M. Weber. 2019. Effects of Background Music on Risk-Taking and General Player Experience. In *Proc of the Annual Symp. on Computer-Human Interaction in Play*. ACM, New York, NY, USA. https://doi.org/10.1145/3311350.3347158

[36] Katja Rogers, Matthias Jörg, and Michael Weber. 2019. Effects of Background Music on Risk-Taking and General Player Experience. In *Proc. of the Annual Symp. on Computer-Human Interaction in Play (CHI PLAY '19)*. ACM, New York, NY, USA. https://doi.org/10.1145/3311350.3347158

[37] K. Rogers, M. Milo, M. Weber, and L. E. Nacke. 2020. The Potential Disconnect between Time Perception and Immersion: Effects of Music on VR Player Experience. In *Proc of the Annual Symp. on Computer-Human Interaction in Play*. ACM, New York, NY, USA. https://doi.org/10.1145/3410404.3414246

[38] Katja Rogers, Giovanni Ribeiro, Rina R. Wehbe, Michael Weber, and Lennart E. Nacke. 2018. *Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173902

[39] Katja Rogers and Michael Weber. 2019. Audio Habits and Motivations in Video Game Players. In *Proc. of the 14th International Audio Mostly Conference: A Journey in Sound*. ACM, New York, NY, USA. https://doi.org/10.1145/3356590.3356599

[40] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task and Environment. In *Proc. of the Annual Symp. on Computer-Human Interaction in Play (CHI PLAY '19)*. ACM, New York, NY, USA, 199–211. https://doi.org/10.1145/3311350.3347190

[41] Timothy Sanders and Paul Cairns. 2010. Time perception, immersion and music in videogames. *Proceedings of HCI 2010 24* (2010), 160–167.

[42] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (2001), 266–281. https://doi.org/10.1162/105474601300343603

[43] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[44] Anthony Steed, Tuukka M Takala, Daniel Archer, Wallace Lages, and Robert W Lindeman. 2021. Directions for 3D User Interface Research from Consumer VR Games. *IEEE Transactions on Visualization and Computer Graphics* (2021).

[45] Zachary O Toups, Andruid Kerne, and William Hamilton. 2009. Motivating play through score. In *In ACM CHI Workshop on Engagement by Design*. Citeseer.

[46] J. von Willich, M. Funk, F. Müller, K. Marky, J. Riemann, and M. Mühlhäuser. 2019. You Invaded My Tracking Space! Using Augmented Virtuality for Spotting Passersby in Room-Scale Virtual Reality. In *Proc. of the 2019 on Designing Interactive Systems Conference*. ACM, New York, NY, USA. http://doi.acm.org/10.1145/3322276.3322334

[47] Julie R. Williamson, Mark McGill, and Khari Outram. 2019. PlaneVR: Social Acceptability of Virtual Reality for Aeroplane Passengers. In *Proc. of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA. https://doi.org/10.1145/3290605.3300310