Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality

Katja Rogers[*], **Giovanni Ribeiro**[**], **Rina Wehbe**[**], **Michael Weber**[*], **Lennart E. Nacke**[**]
[*]Institute of Media Informatics, Ulm University, Ulm, Germany
[**]HCI Games Group, Games Institute, University of Waterloo, Waterloo, Canada
[*]<firstname>.<lastname>@uni-ulm.de, [**]ggsribei@uwaterloo.ca, rina.wehbe@uwaterloo.ca, lennart.nacke@acm.org

ABSTRACT
Sound and virtual reality (VR) are two important output modalities for creating an immersive player experience (PX). While prior research suggests that sounds might contribute to a more immersive experience in games played on screens and mobile displays, there is not yet evidence of these effects of sound on PX in VR. To address this, we conducted a within-subjects experiment using a commercial horror-adventure game to study the effects of a VR and monitor-display version of the same game on PX. Subsequently, we explored, in a between-subjects study, the effects of audio dimensionality on PX in VR. Results indicate that audio has a more implicit influence on PX in VR because of the impact of the overall sensory experience and that audio dimensionality in VR may not be a significant factor contributing to PX. Based on our findings and observations, we provide five design guidelines for VR games.

ACM Classification Keywords
H.5.1. Information Interfaces and Presentation (e.g. HCI): Multimedia Information Systems

Author Keywords
games; audio; virtual reality; player experience; ambient noises; background music; sound effects.

INTRODUCTION
Since *Facebook* bought *Oculus* for $2 billion dollars in 2014, virtual reality (VR) has been centre-stage in commercial applications and research because of its promise to provide a highly immersive audio-visual experience [6]. While researchers have long seen the potential of VR technology for educational applications [7, 12] and therapeutic benefits [46, 33], they have only limited knowledge about what facilitates these immersive VR experiences, and specifically how audio contributes to these impressions. Thus, we still lack a complete scientific understanding of player experience (PX) in VR games, which are fast becoming a lucrative mass-market.

Audio is another important output modality linked to immersion and presence [19, 43] in digital games. However, it has not been scrutinized in the specific context of VR games. Game sound can enhance the realism of the gameworld and induce emotions [3, 18]. Furthermore, game audio also plays a significant technical role in providing basic user feedback, by providing confirmation of player actions or warnings of in-game events [30, 45, 28]. In addition, game developers often have to make budget decisions about the dimensionality and cost of their games' audio design. While previous research has investigated the use of individual audio types, there are no guidelines for game developers about how the inclusion of multiple audio types affects PX in their games.

A meta-analysis of 115 effect sizes found a medium-sized effect of sound quality on presence [15]. Thus, audio has been shown to have a larger effect on presence than image quality. The same analysis showed that the degree of technological immersion is even more influential in developing presence. Despite this study, there remains a lack of research on the effects of audio on immersive experiences in modern head-mounted-display (HMD)-based VR. In particular, researchers have not yet explored how audio in VR contributes to engagement, enjoyment, or to the affective state of players. This led us to pose our general research question; its answer contributes to our understanding of human experiences in VR and helps developers create more immersive VR experiences: *Does audio play a key role as part of the multi-sensory, immersive VR experience or does it not affect this experience?*

To address this research question, we studied the immersive effects of VR as PX and audio perception in a commercial horror-adventure game (*The Vanishing of Ethan Carter* [1] by The Astronauts, 2014) known for its atmospheric sound design [29]. We first conducted a mixed-methods experiment (N=12) by directly comparing the impact of the HMD-VR and monitor-display versions of the same game on survey scales and with a qualitative semi-structured interview.

Our results from this initial study show that there were no significant differences in immersion as operationalized by the Immersive Experience Questionnaire (IEQ). However, the monitor-display condition was perceived as more challenging

DOI: https://doi.org/10.1145/3173574.3173902

while the HMD-VR condition showed more emotional involvement. Our semi-structured interview supported those results, suggesting that audio takes a less significant role for most players in VR, because they focus more on the sensory experience and exploratory gameplay. In comparison, monitor-display players are more concerned with game progression.

To follow up on this, we conducted a second quantitative between-subjects study in VR (N=40) to explore the effects of audio dimensionality, meaning the number of audio types provided using four conditions[1]. By selecting (1) narration and sound effects as the minimal viable game version (i.e., providing critical feedback, our first condition), we then added two more conditions with (2) ambient noises or (3) background music to investigate those factors in more detail, and finally, the (4) "full game" experience with all four audio types. Based on previous research suggesting more sensory stimuli to predict more presence [39], we expected an increase in presence and engagement for a full-dimensional audio design.

Our results indicated that PX is not significantly improved through greater audio dimensionality in VR, which contradicts the above indications from the related literature (but supports other PX literature for monitor-display games [47, 43]). The affective factors of PX operationalized by the Self-Assessment Manikin (SAM) and Positive-Affect-Negative-Affect-Scale (PANAS) were not significant predictors of the audio conditions using a multinomial logistic regression. Immersion via IEQ and engagement via game engagement questionnaire (GEngQ) similarly lacked significant predictive power for the audio conditions. Audio condition did not influence player identification in the game either. Given this multidimensional PX analysis, we can confidently say that audio dimensions likely only have minimal influence on PX in VR.

We make three important contributions with this research. First, we show that aesthetic factors of the gameworld matter more for HMD-VR environments while game progression is more important in monitor-display-based games. Second, we show that audio dimensions do not have a different influence on PX as operationalized by several validated survey scales and that minimal audio does not change PX in HMD-VR environments. Third, we contribute to the growing VR literature in human-computer interaction (HCI) by proposing five design guidelines for VR games: (1) limit audio dimensions in VR games; (2) avoid forced camera movements; (3) limit play session time for gamepad-based VR navigation; (4) design for exploration in VR; (5) accommodate for novelty bias.

Overall, our studies provide a useful exploration of PX in VR. Our results add new contrasting evidence to VR HCI studies about the importance of audio in HMD-VR environments. Finally, our guidelines will help developers implement successful interactive VR entertainment experiences.

RELATED WORK

Two main topics form the subjects of the studies reported in this paper. First, we discuss audio in digital games, and second, we explore the effects of modern VR technology on PX.

[1] In the literature, audio dimensionality has also referred to spatialized audio.

Audio in Digital Games

The literature indicates that audio is an important factor for inducing *presence* in games [45, 35], particularly through sound quality and spatialization [39]. Cummings & Bailenson's meta-analysis reported a medium-sized effect of sound quality on presence [15]. Audio in games has been discussed for its immersive attributes, because it helps to elicit feelings of presence and engagement (i.e., it supports dissociation and transportability). In this paper, we follow Jennett et al.'s [27] definition of immersion in the game experience, which involves a loss of awareness of time and surroundings. Inappropriate, unrealistic, or absent audio can hinder this effect [19]. Particularly in interactive digital games, sound supports cognitive appraisal and immersion by hiding the medium, supporting realism, and strengthening emotional reactions [18].

Game audio has been shown to affect PX through players' engagement and affective states [43, 45]. However, the diversity of digital games has made it difficult to explore the effects of audio in general terms: there are distinct genre differences in the importance and functions of sounds [30]. While audio is rarely used as a game input modality, it can be crucial for expert gameplay through the information conveyed by audio output: For example, by supporting opponent localization [28].

There are many different classifications of audio in games. Liljedahl distinguishes between speech and dialogue, sound effects, and music, with sound effects further classified as ambient sounds, avatar sounds, object sounds, character sounds, and ornamental sounds [36]. Jorgensen [31] and Ekman [17] have instead discussed and categorized game sounds in how they relate to diegesis: whether the audio originates from in-game sources or outside of it, and whether it signifies in-game events or outside events. Nacke et al. also studied diegetic 'sound' (i.e., "sound that originates from the game environment, its objects and its characters") and non-diegetic 'music' (i.e., "the non-diegetic musical soundtrack") in a first-person shooter game, suggesting that sound without the music facilitated 'Flow' in the game [44]. For this paper, we follow Liljedahl's [36] categories to discriminate between speech and dialogue, sound effects, and music. Later, we will distinguish between types of sound effects for the study's game stimuli.

Functions of Audio Types in Games

Different types of audio embody different functions in games. For example, speech and dialogue is used for storytelling and character development. Voice-over added to written dialogue increases player engagement [10]. Interestingly, some studies are now pointing to no significant impact of sound cues and game audio on PX or physiological measures [47, 43].

Listening to music induces emotional responses, both on its own and when accompanying visual media [24, 42]. Thus, it is unsurprising that while background music is often considered peripheral, it is crucial for dramatic storytelling and affects game perception [3, 38]. Lack of background music affects players' perceived control, presence, realism, and attention [30]. Background music also increases immersive qualities of games through its presence, and through dynamic changes [23, 54]. Non-diegetic music was also found to increase feelings of tension in a first-person shooter game when listened to during gameplay without feedback sounds [44]. Interestingly, effects of background music on cognitive functions such as fact recall can depend on the display system [20].

Sound effects are crucial for giving feedback to players about their own in-game actions, and are often used to convey diegetic information, for example informing players of opponents' actions [31, 17, 45] or providing clues about player health [47]. In first-person shooters, sound effects also influence immersive qualities of the game, by facilitating both challenge-based and sensory-based immersion [25]. Proprioceptive sound effects are thought by some to facilitate imaginative immersion and increase player identification, but this has not yet been substantiated [25]. There is further research on ambient environmental sound effects in particular. Ambient noises and sound effects in general affect time perception, as well as player performance and behaviour [11]. Ambient noises increase presence in both real-life and digital environments [35, 9]. This effect has also been demonstrated in VR [16]. Cummings & Bailenson's meta-analysis has suggested that presence is facilitated through self-location and spatial models of the virtual space; thus, environmental sounds may be more important for inducing presence.

Although most games use multiple audio types, combinations of audio types (i.e., high vs. low audio dimensionality) in games are rarely investigated. Lombard & Ditton's comprehensive review has suggested that a higher number of sensory stimuli leads to greater presence [39]. This has not been empirically investigated for game audio types in VR.

Effects of VR on PX

Modern VR technology with head-mounted displays (HMDs) facilitates highly immersive experiences [6]. Side effects can include motion sickness or cybersickness in the form of nausea and dizziness, likely because of perceptual mismatches between users' visual and vestibular senses, although its exact causes and the involved factors are still being researched [2]. Nevertheless, with increasingly sophisticated HMDs and motion tracking, modern VR is becoming more widespread as consumer technology, including use in VR arcades [14].

Presence has been linked to performance and learning [53]; thus, immersive technologies are often investigated for use in education. The growing societal interest in VR has accordingly been accompanied by a diverse range of research on VR's potential application in therapeutic [5] or educational use cases (e.g., language and cultural education, mathematics, and medicine [12, 41]). The results show that VR can be effective for learning, but not significantly better than traditional (monitor-display) digital experiences [12].

Research on the effects of audio within VR is sparse. Early-stage VR showed that the presence of ambient noises and sound effects was able to increase the immersive characteristics of VR, while increased visual detail did not improve it significantly [16]. This suggests that there may be a saturation effect for the visual senses, after which more detail or higher resolution for a particular sense may not improve feelings of presence. To our knowledge, it has not been studied whether this also occurs for non-visual senses.

A majority of studies focus on educational or training games, but Lin investigated fear reactions and coping strategies in participants' responses to a VR horror game [37]. They report greater fear reactions for event-related game elements (e.g., monster jump scares) than environment-related game elements (e.g., darkness, environmental sounds). In her user study, 17% attributed fear reactions to sound effects.

Another strongly related study by Tan et al. [49] compares game experiences in the HMD-VR and monitor-display versions of *Half-Life 2*, using the Oculus Rift HMD. They found the VR experience was more intense, despite feeling less in control. Some actions, such as looking around or aiming, were particularly enjoyable to participants. Thus, they suggest designing for passive game elements. Their paper includes minor mentions of audio perception: one participant found the monitor-display version less engaging, but reported that this effect was lessened by music, while another described an in-game event as sounding more 'more immersive' in VR. However, audio perception was not a focus of their study.

STUDY 1: HMD-VR VS. MONITOR-DISPLAY

To explore the immersive effect of VR technology and audio as a mediating factor, we conducted a mixed-methods study comparing the *HMD-VR* and *monitor-display* versions of the horror adventure game *The Vanishing of Ethan Carter* (*TVoEC*). We chose this genre because it aims to elicit player affect of strong intensity (i.e., fear and anxiety) and sound is considered an important factor mediating this experience [40, 34]. Thus, horror games are often used as stimuli for game audio research (e.g., [44, 26, 22]). We compared effects of the two game versions with validated questionnaires and a semi-structured interview, particularly investigating audio perception.

Participants

We recruited 12 participants (6 female, age M=23.5, SD=4.62). The participants were evenly divided by gender in the counterbalanced *HMD-VR* and *monitor-display* conditions. More than half the participants reportedly played digital games at least every month, 25% played digital games every day; no participants played less often than a few times a year. All but one participant reported that they had never played the game before in any version. Participants indicated that they did not have much experience with VR (M=3.80, SD=1.99 on a seven-point scale), but tended towards expressing an interest in VR (M=5.45, SD=1.88). Almost half (five participants, 41.67%) had never used VR, while four (33.33%) had played a few times. A single participant used VR frequently and also developed for VR. Our participants reflect a standard population for whom VR is not yet a commonly-used technology.

Methods

We conducted a semi-structured interview with the participants to explore their experiences of both game versions, and solicit their perception of the audio (as a whole, and in both versions). The first half of the interview focused on participants' impression of the game in general, as well as their likes and dislikes of both the *HMD-VR* and *monitor-display* conditions (e.g., "*How would you describe something you liked/disliked about the PC/VR version?*"). The second half targeted the audio aspects, by asking about each audio type specifically (e.g., "*How would you describe the narration/sound effects/ambient noises/background music?*" and "*To what degree did you notice it as part of your game experience?*"). Finally, the interview asked about differences in their perception of the audio between the two versions ("*Did your perception of the audio differ between the two versions?*").

The interview transcriptions were analyzed using an iterative thematic analysis approach with three raters [21]. A first codebook was built by all raters independently by coding the first 3 interviews. Raters used an open coding scheme to tag parts of the transcript that were related to the research question. Following this process, the independent raters discussed the emergent themes and tags used. Upon agreement, the codebooks were then combined into one operational set. Following the meeting, reviewer 1 applied the new operationalized code set to the previously coded data. This codebook was then adapted for independent coding of the next three interviews. Next, coders again met to finalize the coding scheme and discussed emerging issues. Finally, the second half of the interview data was coded with the resulting codebook (five new codes were added during that time). Following the analysis of inter-rater reliability scores, which originally were calculated at a maximum of 50% agreement, the data set was discussed and recoded. We calculated inter-rater reliability as a frequency of codes, that is, each code group frequency was divided by the overall number of codes. We calculated agreement while setting our accepted error rate as within 0.05. Upon averaging the score, reliability of all three raters reached 90%.

Further, we used three validated PX questionnaires. The *GEngQ* [8] was used to measure game engagement, consisting of 19 items on a seven-point Likert scale. Positive and negative affect was measured via *PANAS* [52] (20 items for each affect on a five-point scale for multiple adjectives describing player emotions). The experience's degree of immersion was measured with the *IEQ* [27]: a single-item measure of immersion (ten-point scale), and five factors (*challenge*, *control*, *real-world dissociation*, *cognitive involvement*, and *emotional involvement*) through 30 items on a seven-point Likert scale.

Stimuli

Participants were asked to play the first ten minutes of *TVoEC*, once in the *HMD-VR* version, and once using the *monitor-display* version (in counterbalanced order). The game was chosen as it offers both an *HMD-VR* and *monitor-display* version, well-crafted audio, an absorbing narrative arc, and exposure to multiple audio types [29, 32, 13]—which can be turned on/off individually. As is common to many VR games, it was ported from a non-VR game, and its game design does not leverage spatially realistic binaural audio. Further, its first ten minutes are easy for novices to games and VR.

The game opens as player character Nick Prospero begins to investigate the disappearance of Ethan Carter. The main game mechanic in the segment consists of using a bubble (which appears upon inspecting objects) to show clues. With each clue, the bubble size increases. In the first segment, players walk along rail-car tracks through a forest (traps therein enlarge the bubble). After the forest, players arrive at a bridge with a broken down rail-car, followed by a murder scene. If players inspect enough objects, the bubble mechanic leads them to find a crank that fixes the rail-car.

As mentioned, we follow Liljedahl's classification of audio types (speech and dialogue, sound effects, music) [36]. We further distinguish ambient and UI sound effects, based on Lin's findings that players have different fear reactions to environment-related game elements [37] and previous research suggesting that facilitating spatial understanding of a VR environment increases presence [15]. With this game stimulus, participants were exposed to multiple types of audio: voice/narration (with subtitles), sound effects (e.g., the player character's footsteps, and feedback sounds for UI elements), ambient noises (e.g., wind rustling through grass and canopy, birds), and background music. The game was played with a VR-HMD (HTC Vive) or 27" ViewSonic VX2703 monitor display depending on the condition, a Microsoft XBox 360 controller, and Sennheiser HD 205 over-ear headphones.

Procedure

After an introduction to the study procedure, written consent was obtained. Participants were then given a questionnaire on their demographic background and their prior experience with the stimulus game. In lieu of a practice session, participants were given a written description of the main game mechanic (finding clues with the bubble mechanic).

The study then proceeded as follows: (1) Each participant received a written introduction to the controls. For the *HMD-VR* condition, we reminded participants to inform the student researcher if they started to experience any simulator sickness. (2) Participants were asked to play *TVoEC* in either the *HMD-VR* or *monitor-display* condition for 10 minutes. (3) They were asked to fill in the PX surveys for that condition. This steps-sequence (introduction to controls, gameplay, and standardized surveys) was repeated once more for the remaining version (*HMD-VR* vs. *monitor-display*).

Finally, we asked participants to take part in the audio-recorded semi-structured interview. Afterwards, we thanked them for their time, entered them in a draw for a $20 Amazon voucher, and offered them candy and a feedback form.

Note on Simulator Sickness

Participants who experienced symptoms were offered breaks from playing, or allowed to withdraw entirely. Overall, simulator sickness was not a pervasive problem in this study: Three participants asked to take a brief break during the VR condition because of motion sickness. In these cases, they were invited to sit down and offered water. After approximately 1 minute, they resumed playing. A fourth participant reported slight vertigo after the study, but did not pause the game.

Results

The study yielded a quantitative comparison of PX components through the questionnaires, and a qualitative comparison based on the transcriptions of the audio-recorded interviews.

Quantitative Results

Overall, game engagement resulted in a positive mean (M=4.04, SD=1.08). Positive affect was low (M=2.99, *SD*=0.74), but higher than negative affect (*M*=1.68, *SD*=0.61). Comparing between conditions, game engagement (t=(7)=0.32, *ns.*[2]) yielded no significant differences; neither did positive (t=(11)=0.45, *ns.*) and negative (V=39, *ns.*) affect. However, game engagement in the *HMD-VR* condition showed a significant negative correlation with players' rating of their previous VR experience, r_{τ}=-0.56, $p<0.05$.

Surprisingly, there was no significant difference in immersion as measured by the IEQ single-question measure (V=29, *ns.*) or the IEQ overall score (V=56.5, *ns.*). Only two factors within the IEQ showed a significant difference: a dependent t-test showed that the *monitor-display* condition (*M*=4.35, *SD*=1.21) was perceived as significantly more challenging than the *HMD-VR* condition (*M*=3.94, *SD*=0.89), t=(11)=-2.22, $p<0.05$ (see Fig. 1). The second significant difference was found in the emotional involvement factor, which was significantly higher in the *HMD-VR* condition (*M*=5.31, *SD*=1.15) than in the *monitor-display* condition (*M*=4.69, *SD*=0.95), t=(11)=2.40, $p<0.05$. Scores for cognitive involvement (*M*=5.36, *SD*=0.69) and control (*M*=5.27, *SD*=0.98), and real-world dissociation (*M*=4.76, *SD*=1.09) did not differ significantly. There was no correlation between players' previous VR experience and their emotional involvement or their challenge scores.

Qualitative Findings

In the qualitative data, several themes and differences emerged from a close analysis of the interview transcriptions.

Difference in Immersion. In contrast to the quantitative results, the interviews revealed that the majority of participants clearly perceived the *HMD-VR* version to be more immersive than the *monitor-display* version: "*felt more immersive*" (P7, *VR→PC*), "*like you're more into the game*" (P9, *VR→PC*), "*felt like I was there*" (P5, *VR→PC*). While talking about the *HMD-VR* version, participants often mentioned its novelty: "*it's just fresh and it's cool*" (P2, *PC→VR*), "*when I was playing on the computer I was thinking, "If this is how I feel [...] being within the virtual world, it will be [...] more exciting." I was really excit[ed]*" (P4, *PC→VR*).

Distraction Factors. When participants stated a preference for the *monitor-display* version, this often coincided with mentions of motion sickness or slight discomfort: "*I felt like throwing up [...] I felt dizzy*" (P6, *PC→VR*), "*feeling like I was going to trip on something*" (P8, *PC→VR*), "*when the game was moving me [...] I felt a little uneasy [...] I wasn't in control of that movement, and it made me feel more nauseous*" (P12, *PC→VR*). In contrast, in the *monitor-display* version different distractions made participants aware of their surroundings and the game setup: "*you turn and there's a wall, you're not fully in it, like in VR*" (P9, *VR→PC*), "*I was just concentrated on the screen*" (P4, *PC→VR*), "*I could see the computer screen and I was actually in the real world, but in the VR I felt like I was actually in there*" (P5, *PC→VR*).

Difference in Gameplay Absorption. When asked specifically in which version participants felt more absorbed in the gameplay, both versions were mentioned, which may account for the lack of significant differences in the quantitative data. However, more detailed questions revealed nuances of this absorption. In describing the *HMD-VR* condition, participants focused on the visuals and the feeling of being present in the gameworld: "*more focused on seeing what was around me*" (P8, *PC→VR*), "*more focused on physically looking around [...] I was moving my head around, so I think was focusing on the vision and what I'm trying to see*" (P5, *VR→PC*). They also indicated a greater focus on the sensory experience in VR: "*more focused on my senses experience, my whole body experience*" (P4, *PC→VR*), and "*paying too much attention to my motions*" (P10, *PC→VR*). Further, they described moments of exploration ("*more focused on getting places*" (P7, *VR→PC*)), and pausing to appreciate the view (e.g., "*Looking at the [sky] and seeing [...] the wind was blowing around me.*" (P4, *PC→VR*)). In contrast, they described an increased focus on the gameplay progression for *monitor-display*. They were more focused on finding clues: "*focus more on the game because [...] I found a lot more clues*" (P5, *VR→PC*), and "*too busy trying to [...] win*" (P9, *VR→PC*). For some, this translated into greater immersion ("*more immersed in playing the game, getting it done, and the puzzles*" (P11, *VR→PC*)), while for others, "*it felt more immersive in VR, and PC [...] felt like a task.*" (P7, *VR→PC*).

Functions of Audio Types. When asked about the different audio types, participants' experiences clearly differed in which aspects they considered important and how they responded to them. Background music and ambient noises were most often considered to hold the most importance. For some, the background music had the greatest effect on their experience, while for others, the ambient noises had greater impact. Interestingly, half of the participants did not notice the presence of background music at all: "*didn't notice music at all*" (P10, *PC→VR*), "*There was not any*" (P11, *VR→PC*), or "*There was music?*" (P12, *PC→VR*).

When it was consciously perceived, background music was often mentioned in combination with emotional responses: "*on edge*" (P6, *PC→VR*), "*really calming*" (P2, *PC→VR*), "*affected my mood*" (P5, *VR→PC*). This was often mentioned in the context of reactions to adaptively changing music. For example, "*a little bit more frightening when you went off the rails*" (P6, *PC→VR*). The adaptive music influenced players' expectations ("*when the music [changes] tone, I expect [...] something jumping out at me*" (P1, *VR→PC*)) and their gameplay: "*the music changed [...] more action-based music. So, I knew okay, I have to follow him*" (P2, *PC→VR*), or "*the sound [...] guide[d] me through more actions*" (P4, *PC→VR*).

Ambient noises tended to be mentioned related to realism and presence: "*realistic*" (P6, *PC→VR*), "*added to [...] depth of how you felt there [...] looked down at the grass, and you saw it blowing and you like heard it*" (P9, *VR→PC*). Some participants specifically brought up changes in the ambient noises: "*When you [...] stepped, [...] it would [...] crunch, with [...] leaves or [...] if you were on the wood [...] it would [...] change.*" (P9, *VR→PC*). One participant described feeling more relaxed because of the ambient noises: "*like peace [...] de-stressed*" (P4, *PC→VR*).

[2] *ns.*: not significant, i.e., $p > 0.05$

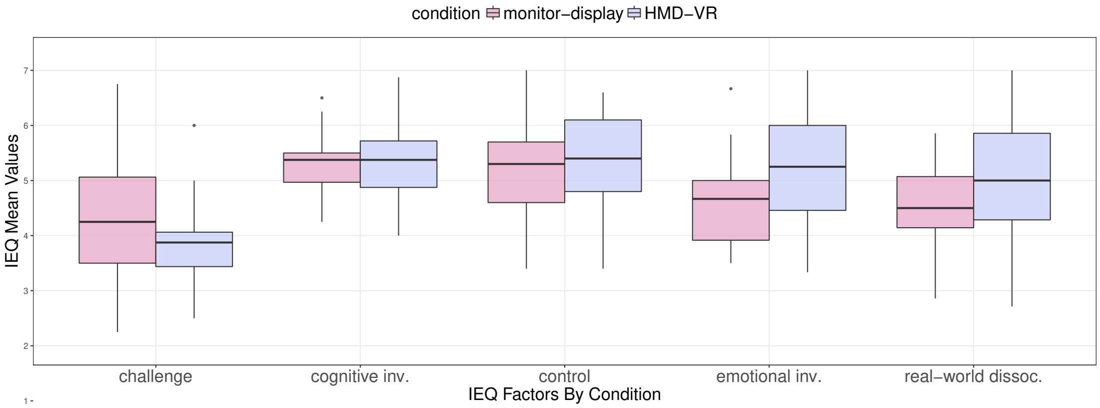

Figure 1. Factors of IEQ for the study comparing *monitor-display* version of the game to the *HMD-VR* version. Emotional involvement was significantly higher for *HMD-VR*, while challenge was significantly higher for *monitor-display* (p<0.05).

Narration was deemed less important by most participants: "*the narration didn't affect [me] much because [...] I didn't care*" (P4, PC→VR). Most agreed that it was necessary and was "appropriate" for the setting, for example "*that was good too, creepy [...] he was [...] scaring me*" (P6, PC→VR).

Sound effects similarly received less attention, for example "*nothing really stood out to me as in the wrong place*" (P8, PC→VR). However, they were mentioned as catching and guiding attention: "*[...] I didn't see it, but then I heard the sound, it was so sudden that I was scared a bit*" (P2, PC→VR).

Difference in Audio Perception. Participants reported differences in the way they perceived the game audio in the two versions. For the majority of participants, audio shifted into the background in VR: "*I paid more attention to the sounds of the game when it was on PC, versus not so much when it was in the VR [...] background noise where you didn't really pay attention to it*" (P8, PC→VR). It was considered more noticeable for the *monitor-display* condition: "*more focused on the listening part while I'm playing with the PC [...] the background music was more triggering*" (P5, VR→PC), "*the audio aspect affected me more in the PC game*" (P7, VR→PC), This phenomenon was mentioned for audio in general, as well as for each audio type specifically: "*the narrative was stronger on PC [...] of course, the audio's the same, but it felt stronger on PC version*" (P2, PC→VR), "*but in the VR version, I didn't notice [the music] at all*" (P10, PC→VR). For some, this manifested as a difference in volume perception: "*it felt louder on the desktop version*" (P11, VR→PC), "*when I was playing the VR version, it was a little quieter*" (P12, PC→VR). The volume was, of course, calibrated equally in both conditions.

Participants attributed this difference in audio perception to the sensory experience of VR: "*I was still [...] wrapped up with the whole VR experience. I [...] focused more on where I was walking rather than what was going on in my ears*" (P7, PC→VR). In particular, this was tied to distractions from the visual senses ("*focusing on the vision*" (P5, PC→VR)), as well as the bodily senses: "*paying too much attention to my motions, maybe that's why*" (P10, PC→VR).

Study 1: Discussion

The study showed few differences between *HMD-VR* and *monitor-display* conditions for engagement and player affect in the quantitative results. The *PANAS* questionnaire may not have been the best instrument for measuring this particular game segment. The emotional responses may be too subtle for the adjectives used in the *PANAS* measure, because the stimulus is exploration-based in nature and only offers a few traps which may startle the player. The active segments later in the game might elicit a greater distinction by causing more intense emotional responses. However, this may also imply that the use of VR did not immediately increase engagement or influence players' affective state. In the comparative study by Tan et al. [49], specific game interactions involving vertical navigation or aiming the gun were considered particularly enjoyable in VR. Full-body movement has been shown to increase engagement and intensify affective experiences [4]. Therefore, *TVoEC* may simply not offer enough physicality in its first ten minutes to elicit similar responses in VR.

The lack of significant differences of player immersion is perhaps more surprising. VR has high technological immersion while regular desktop computers have low technological immersion; the meta-analysis of Cummings & Bailenson reported a medium-sized effect of such high-vs.-low manipulations on presence [15]. However, the *IEQ* was designed to explore immersion as related to but not equal to presence; it is thus more closely related to game engagement (*GEngQ*) which also did not differ based on the display version.

However, the game version did influence participants' emotional involvement, which was significantly higher in the *HMD-VR* condition. In comparison, playing in the *monitor-display* condition elicited a significantly greater sense of challenge. This is interesting, and can be further explored through the analysis of the qualitative data. The interviews revealed a difference in players' style of gameplay depending on the condition they played in. One of the themes that emerged from the interviews saw that participants were focused more on the gameplay mechanics and their progress in the game (i.e., finding clues to solve the mystery) when playing on the monitor-display than when they were in VR. This may explain why the *monitor-display* version was perceived as significantly more challenging; the players were more focused on the game task and the gameplay progression.

Inversely, the *HMD-VR* condition provided a more sensory-immersive experience. Players were more absorbed by the sensory experience and the visuals (despite being a worse resolution). This difference in VR gameplay focus may explain the significantly higher means for emotional involvement. Their more indulgent exploration of the gameworld and the sensory experience may have facilitated a higher emotional involvement in the *HMD-VR* condition. Some of these results were already hinted at in Tan et al.'s findings: they report that players particularly enjoyed looking around in VR and suggest designing for passive game elements. The difference in gameplay style found in our qualitative results contributes to a better understanding of their results.

This difference in gameplay style may, of course, be influenced by novelty bias. Participants could have been more distracted from the game task by the sensory experience of being in VR. While playing on the monitor-display, they are generally familiar with the setting, and can thus focus more on the gameplay. Regardless of the factors involved, this is a phenomenon that must be considered by current research on and development of VR games, as this lack of experience with VR is typical of many, if not most, people in the current population. This is supported by the inverse correlation of prior VR experience with game engagement in our study.

The game segment contained sequences, during which participants indicated discomfort relating to motion sickness; notably at the beginning when they experience a forced camera movement along the train tracks and later a pile of virtual stones that many participants walked over rather than around. Here, sensory mismatches can translate to dizziness or nausea [2].

Difference in Audio Perception

Participants reported a clear difference in perception of the audio depending on the game version, meaning audio was perceived as less noticeable or less significant to their experience in VR. We suggest that a combination of novelty bias and sensory distraction (through the visual senses, the vestibular system, and potential motion sickness) is a strong component of this phenomenon. Until commercial VR technology becomes more widely used, it can only be speculated how much of the previously mentioned difference (in how audio was perceived) is caused by the VR experience itself, rather than influenced by the novelty or motion sickness factors that accompany it. A study targeting expert long-term users may be able to explore this question in the future.

Overall, we must acknowledge that participants may also be self-selecting, because participants who know that they are susceptible to motion sickness may be less likely to volunteer to participate. Additionally, some people were interested in participating but withdrew when they learned that the game was classified as horror adventure (although the segment used as the stimulus is not typical of this genre). The remaining participants may therefore have self-selected with a preference towards the horror genre, which correlates with personality traits such as sensation-seeking behaviour [40].

Finally, we acknowledge potential genre effects. The role and function of audio in games played on the PC has been shown to differ based on the game genre [30, 25]; this may also apply to VR games. For example, the use of a more action-filled game segment may see the role of audio become more prominent in VR (e.g., a first-person shooter (FPS) as also used by Tan et al. [49], given that sound in FPS games is important for challenge-based immersion [25]).

STUDY 2: AUDIO DIMENSIONALITY

The first study indicated that audio as a whole may be perceived less prominently in VR. As such, we were interested in exploring the effects of the sensory richness of audio in VR, that is, audio dimensionality. Our emerging research question was: *If audio is less important in VR, does it matter how many audio types are provided?* Sensory resolution or richness has been shown to increase presence in virtual environments [50, 48, 39, 53]. However, for the impact of visual senses, there have been indications of a potential saturation effect, that is, after a certain degree of visual richness is provided, you may not make much difference to the visual experience by adding more vividness. [16]. Therefore, our next study focused on the importance of audio dimensionality for player experience.

We conducted a quantitative between-subjects study to investigate effects of different numbers of audio types on PX in the *HMD-VR* version of the game. As the game becomes unusable without at least narration to introduce the plot, and sound effects to help the player interact via the game mechanics, we decided against using a condition without any audio at all. Instead, we chose our version of minimal audio dimensionality to consist of only narration/voice (*V*) and (non-environmental) sound effects (*S*). To isolate effects of environmental sound effects on the immersive experience (based on related work indicating effects of spatial information [15]), we then added a condition that extended the minimal version with ambient noises (*A*). To leverage the effects of emotional responses to music, our third condition consisted of the minimal version with background music (*M*). Finally, the fourth condition provided the "full" game audio consisting of narration/voice, sound effects, ambient noises, and music.

Participants

We recruited 40 participants (29 male, 11 female) with an average age of 26.93 (*SD*=4.74). An additional participant was recruited, but their data was discarded as they withdrew from the study because of motion sickness within the first two minutes of the game. The remaining participants were evenly divided into the four groups. Each group had three female participants, except the *VSAM* condition which consisted of two female participants. The participants were all at least casual gamers: Five reported playing a few times a year, and another five played games every month. The majority played either every week (18, 45%) or every day (12, 30%). On average, most participants professed an interest in VR technology (*M*=5.45, *SD*=1.88), but had only little experience with it (*M*=3.8, *SD*=1.99) on a seven-point Likert scale (1=*strongly* *disagree*, 7=*strongly agree*). Three of the participants had played *TVoEC* before, but none had ever played it in VR.

Methods

Like the previous study, this study assessed game engagement (*Game Engagement Questionnaire* [8]), immersion (*IEQ* [27]), and positive and negative affect (*PANAS* [52]). This was extended to include a second measure for players' affective state as *valence*, *arousal*, and *dominance* (*SAM*; each dimension rated on a single semantic picture scale), as the PANAS had proved less reliable in the previous study. Further, related work has suggested that proprioceptive sound effects (i.e., relating to the player character in the form of footsteps or breathing) may facilitate player identification [25]. Thus we included two factors to measure identification with the player character: *similarity identification* and *embodied presence* from Van Looy's player identification scale (*PIS* [51]; each factor is assessed through six items on a five-point Likert scale).

Participants' gameplay was video-recorded as a screen capture. Notes were taken on motion sickness and any comments volunteered by the participants.

Stimuli

In this study, participants played ten minutes of the *HMD-VR* version of the game only. The game offers exposure to and control of the audio types required for our comparison. As described, the participants were assigned to play the *HMD-VR* game in one of four conditions (audio settings were applied before participants were present):

- the full game audio experience: voice, sound effects, ambient noises, and background music (*VSAM*)
- without background music: voice, sound effects, and ambient noises (*VSA*)
- without ambient noises: voice, sound effects, and background music (*VSM*)
- with neither ambient noises nor background music: voice and sound effects only (*VS*)

In the first ten minutes of gameplay, participants were all exposed to the narrative introduction of the game. Specific in-game locations trigger further auditory narrative elements. Sound effects occurred when participants interacted with objects, for example, triggering traps or inspecting items at the murder scene. In the conditions that included this aspect, ambient noises were present at all times, for example, as sounds of wind or birds. Background music was also present throughout the gameplay (in the conditions that included it); however, its prominence depended on the players' location. For instance, upon approaching the bridge, the music picks up both pace and volume to convey greater intensity.

Procedure

Introduction, obtaining consent, collection of demographic data, and remuneration followed the same procedure as the previous study. The participants were asked to play *TVoEC* in only the *HMD-VR* version for ten minutes, and then asked to fill in the standardized surveys on their experience.

Results

In total, sixteen participants experienced motion sickness ranging from slight vertigo to dizziness or nausea. Of these, eight participants asked to take a break during the study. All but one of these felt able to continue after a brief break (approx. 1 min) and water – the final participant took three breaks overall. Of the participants who had played *TVoEC* before, one was assigned to the *VSM* condition; this participant did not notice the missing ambient noises until after the study debriefing.

SAM & PANAS

Overall, participants reported positive values for *valence* (*M*=4.38, *SD*=1.51), *dominance* (*M*=4.65, *SD*=1.56), and *arousal* (*M*=3.80, *SD*=1.47).

Participants' reported valence was highest for the *VSM* (*M*=4.7, *SD*=1.89) and *VSAM* (*M*=4.7, *SD*=1.42) conditions, while results for the *VS* (*M*=4.1, *SD*=1.52) and *VSA* (*M*=4.0, *SD*=1.89) conditions were lower. Arousal was highest for *VSA* (*M*=4.3, *SD*=0.95), and lowest for *VS* (*M*=3.4, *SD*=1.43). Dominance values were highest for *VSM* (*M*=5.2, *SD*=1.48) and lowest for *VSAM* (*M*=4.2, *SD*=1.62). We conducted a multinomial logistic regression using the SAM factors to predict the condition; this yielded no significant predictive power over the null model, $\chi^2(9)$=6.93, *ns.*

Like in the previous study, participants' scores for positive and negative affect were overall low (means ranging between *M*=2.62 and *M*=3.22 for positive, and less than two for negative affect). Positive and negative value were also not significant predictors of the audio condition, $\chi^2(6)$=4.08, *ns.*

IEQ

The means for all IEQ factors were overall high. The means were highest for cognitive involvement (*M*=5.13, *SD*=0.92) and control (*M*=5.09, *SD*=0.94), and lowest for challenge (*M*=4.13, *SD*=1.41). A plot of the IEQ factors compared by condition can be seen in Figure 2. Unlike the previous study, none of IEQ factors showed a significant effect of condition, as shown by a logistic regression, $\chi^2(15)$=17.56, *ns.*

GEngQ

The overall game engagement scores were positive (*M*=3.90, *SD*=0.72). Mean results for *VSAM* (*M*=4.15, *SD*=0.85) and *VSM* (*M*=4.15, *SD*=0.67) conditions were highest and almost identical, followed by *VSA* (*M*=3.66, *SD*=0.64). The *VS* condition showed the lowest mean scores (*M*=3.41, *SD*=0.54). Despite this tendency, a logistic regression showed that there was no significant predictive power of game engagement between conditions, $\chi^2(3)$=5.10, *ns.* (although with p=0.16 it was the closest indication of an effect of audio condition).

Player Identification

Participants' mean results for the similarity identification factor were generally low. Similarity identification yielded a negative mean value (*M*=2.16, *SD*=0.95), while embodied presence was slightly positive (*M*=3.24, *SD*=0.85). A logistic regression showed that neither factor had predictive power over the condition, $\chi^2(6)$=1.63, *ns.*

Overall the results show that audio dimensionality did not have the effect on PX that is suggested by the literature. Instead,

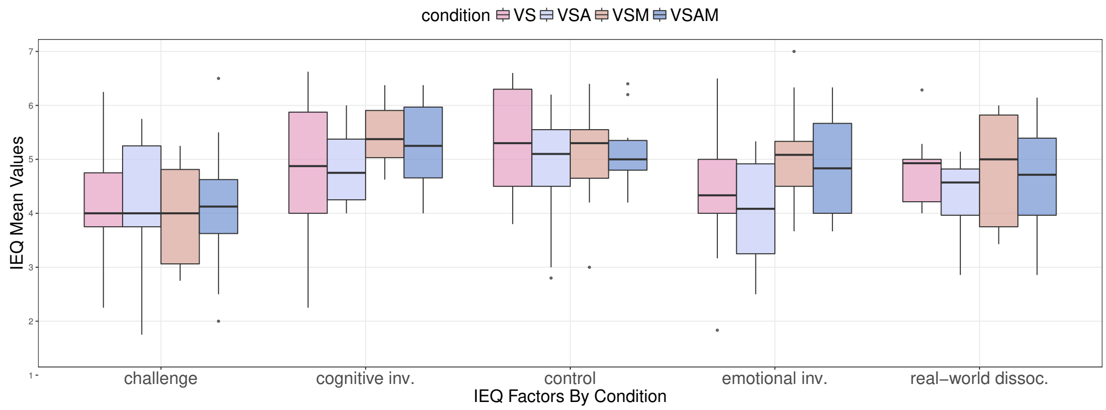

Figure 2. Factors of IEQ compared by audio condition showed no significant differences (p>0.05) for increased audio dimensionality.

there appears to be a saturation effect: adding audio beyond the minimal condition does not significantly impact PX. We will present potential implications for the games user research and HCI community in the discussion.

Post-Study Follow-up: VR Flashback

A few weeks post-study, a participant got in touch with the authors to report a moment of dizziness and feeling "back in the game" while crossing a wooden bridge during a hiking trip. Likely this was triggered by the similarity of the real-world environment with the game layout. To our knowledge, this long-term effect has not been reported in the literature for modern HMDs (for older VR technology, it has been reported as a side effect of presence, cf., [39]). We can only speculate whether other participants might experience the same if confronted with real-life scenarios similar to the VR game.

Study 2: Discussion

The literature suggests that an increased resolution of the sensory stimuli should improve the sense of presence (e.g., [50, 48, 39, 53]). In particular, the addition of ambient noises, with their potential to facilitate players' spatial understanding of the virtual environment, were strongly suggested to mediate the immersive effect of VR. Instead, the results show a surprising lack of effect of audio dimensionality on all measured PX components. Participants were similarly immersed, engaged, and experienced similar affect regardless of whether the game provided narration and sound effects only, or added ambient noises and/or background music. While game engagement showed the clearest trend towards an effect of audio condition, this too proved not to be significant.

The PANAS yielded similar scores as the previous study—given the identical game segment, this consolidates our findings from study 1. We suggest that the questionnaire is not well-suited for game stimuli that are mainly exploration-based and calm in tone. The SAM yielded more useful results; measuring the affective state in terms of valence, arousal, and dominance appears to assess players' emotions more accurately. The SAM scores show a high variance. The previous study suggested that emotional response to audio is a subjective matter; this may have led to these results.

The low values for similarity identification with the player character could be related to the participants only playing for ten minutes; with the game narrative's calm pace, they were not exposed to much character development. Embodied presence showed more positive results. As the minimal version included sound effects such as player steps, it will fall to future work to investigate whether increasing the degree of proprioceptive sound effects can improve factors of player identification further, as suggested by Grimshaw [25].

Similar to the embodied presence, the IEQ results were positive overall. It appears that after a certain degree of sound quality is reached (i.e., the minimal quality necessary for user feedback), increasing resolution through additional audio types does not significantly impact PX in VR. Like Dinh et al. [16] have proposed for the visual senses, we thus suggest there may be a saturation effect for game audio in VR.

Game engagement, while not significant, showed the strongest tendencies toward an effect of audio dimensionality on PX. Drawing on previous evidence of genre differences [30], we suggest that future work focus on investigating game engagement regarding potential effects of audio dimensionality in other game genres. It seems quite possible that attributes of game genre mediate the role of audio in VR games.

Effects of audio dimensionality that the current study was not able to detect might still exist, given the limitations of the low sample size in each group. However, our review of existing literature on audio in traditional games suggested an effect on immersion, because this has been shown with similar sample sizes. Replicating the study with a VR game less prone to causing motion sickness (e.g., room-scale only) might also yield different results, although a previous study suggests that cybersickness does not always impact immersion [49]. Further, there are few VR games with similar well-developed, multi-dimensional sound design.

GENERAL DISCUSSION AND DESIGN GUIDELINES

Together, our two studies suggest that the role of audio may be less pronounced in VR games than it is in traditional PC games. As such, it is perhaps less surprising that there was no significant effect of audio dimensionality in the second study.

Further, the importance placed by players on individual audio types appears to be a highly subjective matter. The interview showed that background music can create emotional responses for some players in VR (particularly through adaptive changes in the music), while others do not even register its presence. Others consider ambient noises to have the greatest effect by conveying realism. This suggests that there may be potential to cater to different audio preferences for players of VR games.

In line with prior research from Robb et al. [47], who suggested design recommendations for informative UI audio in FPS games, we discuss five guidelines for game design and research design concerning immersive VR games that flow from the implications of our study results:

Limit Audio Dimensions in VR Games

The results of the second study suggest a potential saturation effect for audio dimensionality in VR. The importance of audio dimensions may be limited in VR games—particularly in exploration-based adventure games like the one we studied. While basic game feedback sounds [47] are crucial for the creation of a successful VR experience to provide player feedback and facilitate storytelling, other additional audio types do not necessarily impact PX. Future work will explore how this varies depending on audio prominence, which varies by genre.

Avoid Forced Camera Movements

Based on our observations and participant comments, most occurrences of motion sickness were caused by perceptual mismatches stemming from vertical navigation (i.e., traversing the virtual stone pile) and forced camera movements. This was likely included in the game because it had been designed in the same manner in its non-VR original. However, we strongly suggest that developers reconsider their design choices when porting a traditional video game to VR.

Limit Play-Session Time for Gamepad-Based VR Navigation

We suggest that ten minutes is an adequate duration for play sessions in VR research for stimuli that primarily utilize gamepad-based navigation. With ten-minute play sessions, the majority of participants were able to complete the study without pause, and those that did request a break were still willing to continue playing after a brief rest and water. We also suggest that researchers advise participants during recruitment to avoid eating directly before VR studies.

Design for Exploration in VR Gameplay

The participants clearly demonstrated a different focus in their style of gameplay for the *HMD-VR* condition. While they were more progression- and task-focused on the monitor-display, in HMD-VR they showed an increased interest in exploring the gameworld, looking around, and enjoying the scenery. Because of these results, we recommend that VR developers design games and experiences to explicitly accommodate this difference in player interest. This points to a shift in the importance of aesthetics in VR games. Gameplay should accommodate exploration and appeal to visual senses, by allowing for moments of reflection and visual appreciation of the game aesthetics. Future work will have to investigate to what degree this gameplay difference is a mediating factor in causing the difference in how audio was perceived in *HMD-VR* and *monitor-display* versions.

Accommodate for Novelty Bias

Novelty bias and effects of motion sickness are clearly prevalent among VR users in the general population. Future work will have to explore how these impact the differences in gameplay and audio perception. However, until VR becomes more widespread, VR developers will have to design VR experiences with their audience in mind; for the foreseeable future, VR users will largely be novices.

FUTURE WORK

The studies conducted in this work point towards a number of interesting subsequent research questions. For example, does the higher emotional involvement in VR stem from the greater visual sensory focus, as described by participants? Future work will explore whether this finding still occurs in VR games that are less exploration-based, to study how genre differences found for audio in traditional video games [23, 30] apply to VR games. Other future avenues to explore include effects of audio on PX for expert VR users, as well as personalization of audio for players that more highly value background music over ambient noises or vice versa.

CONCLUSION

Our studies in this paper addressed the lack of research on the effects of audio on immersive experiences in VR. We conducted two studies: (1) A within-subjects experiment with 12 participants using a commercial horror-adventure game to study the effects of VR and monitor-display versions of the same game on PX. (2) A between-subjects study of 40 participants on the effects of audio dimensionality on PX in VR. When comparing the *HMD-VR* condition to the *monitor-display* condition, we found that audio played a less prominent role in VR and that participants were more preoccupied with their sensory experiences in VR. In the second study, subjectivity was found in the perceived importance of audio types. Overall, our research has presented three important HCI contributions: (1) We found that visual and aesthetic factors mattered more for participants in the *HMD-VR* condition, while, in the *monitor-display* condition, they were more concerned about the pursuit of in-game progression. (2) We found that audio dimensionality does not influence PX in VR games. (3) We discussed five guidelines which emerged from our studies for developing and researching immersive VR experiences in games. This research presents an important step towards understanding effects of VR game audio on player experience.

ACKNOWLEDGMENTS

We thank all volunteers for feedback on previous document versions, help with recruiting participants, and video footage. We gratefully acknowledge the NSERC Engage grant #075725 supported by Red Meat Games Inc. Lennart Nacke also gratefully acknowledges NSERC, SSHRC, Mitacs, and CFI for funding his research.

REFERENCES

1. The Astronauts. 2014. *The Vanishing of Ethan Carter*. Game [VR, PC]. (26 September 2014).

2. Judy Barrett. 2004. *Side Effects of Virtual Environments: A Review of the Literature*. Technical Report. Command and Control Division Information Sciences Laboratory. Defence Science and Technology Organisation, Canberra, Australia.

3. Axel Berndt and Knut Hartmann. 2008. The Functions of Music in Interactive Media. In *Interactive Storytelling: First Joint International Conference on Interactive Digital Storytelling (ICIDS)*, Vol. 5334. Springer, Berlin/Heidelberg, Germany, p. 126.

4. Nadia Bianchi-Berthouze, Whan Woong Kim, and Darshak Patel. 2007. Does Body Movement Engage You More in Digital Game Play? and Why?. In *Proceedings of the Second International Conference on Affective Computing and Intelligent Interaction (ACII)*, Vol. 4738. Springer Science & Business Media, Berlin/Heidelberg, Germany, p. 102.

5. Cristina Botella, Javier Fernández-Álvarez, Verónica Guillén, Azucena García-Palacios, and Rosa Baños. 2017. Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review. *Current Psychiatry Reports* 19, 7 (2017), p. 42.

6. Doug A Bowman and Ryan P McMahan. 2007. Virtual reality: how much immersion is enough? *Computer* 40, 7 (2007).

7. Meredith Bricken. 1991. Virtual Reality Learning Environments: Potentials and Challenges. *SIGGRAPH Computer Graphics* 25, 3 (July 1991), pp. 178–184. DOI: http://dx.doi.org/10.1145/126640.126657

8. Jeanne H Brockmyer, Christine M Fox, Kathleen A Curtiss, Evan McBroom, Kimberly M Burkhart, and Jacquelyn N Pidruzny. 2009. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology* 45, 4 (2009), pp. 624–634.

9. 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, Berlin/Heidelberg, Germany, pp. 102–111.

10. JaeHwan Byun and Christian S Loh. 2015. Audial engagement: Effects of game sound on learner engagement in digital game-based learning environments. *Computers in Human Behavior* 46 (2015), pp. 129–138.

11. GG Cassidy and RAR MacDonald. 2010. The effects of music on time perception and performance of a driving game. *Scandinavian Journal of Psychology* 51, 6 (2010), pp. 455–464.

12. Alan Cheng, Lei Yang, and Erik Andersen. 2017. Teaching Language and Culture with a Virtual Reality Game. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 541–549. DOI: http://dx.doi.org/10.1145/3025453.3025857

13. Tom Cole, Paul Cairns, and Marco Gillies. 2015. Emotional and Functional Challenge in Core and Avant-garde Games. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15)*. ACM, New York, NY, USA, 121–126. DOI:http://dx.doi.org/10.1145/2793107.2793147

14. Daniel Cooper. 2017. VR arcades need to be social to succeed. Engadget, https://www.engadget.com/2017/06/01/vr-arcades-social-htc-viveland-taiwan/. Accessed: 15th September 2017. (January 2017).

15. James J Cummings and Jeremy N Bailenson. 2016. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology* 19, 2 (2016), pp. 272–309.

16. Huong Q Dinh, Neff Walker, Larry F Hodges, Chang Song, and Akira Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings of Virtual Reality*. IEEE, Piscataway, New Jersey, United States, pp. 222–228.

17. Inger Ekman. 2005. Meaningful noise: Understanding sound effects in computer games. *Proceedings of Digital Arts and Cultures* (2005).

18. Inger Ekman. 2008. Psychologically motivated techniques for emotional sound in computer games. *Proceedings of the Audio Mostly Conference (AM)* (2008), pp. 20–26.

19. Inger Ekman. 2013. On the desire to not kill your players: Rethinking sound in pervasive and mixed reality games. In *Foundations of Digital Games (FDG)*. pp. 142–149.

20. Eric Fassbender, Deborah Richards, Ayse Bilgin, William Forde Thompson, and Wolfgang Heiden. 2012. VirSchool: The effect of background music and immersive display systems on memory for facts learned in an educational virtual environment. *Computers & Education* 58, 1 (2012), pp. 490–500.

21. Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods* 5, 1 (2006), 80–92. DOI: http://dx.doi.org/10.1177/160940690600500107

22. Tom Garner, Mark Grimshaw, and Debbie Abdel Nabi. 2010. A Preliminary Experiment to Assess the Fear Value of Preselected Sound Parameters in a Survival Horror Game. In *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound (AM '10)*. ACM, New York, NY, USA, Article 10, 9 pages. DOI:http://dx.doi.org/10.1145/1859799.1859809

23. Hans-Peter Gasselseder. 2014. Dynamic Music and Immersion in the Action-adventure an Empirical Investigation. In *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound (AM '14)*. ACM, New York, NY, USA, Article 28, 8 pages. DOI: http://dx.doi.org/10.1145/2636879.2636908

24. William W Gaver and George Mandler. 1987. Play it again, Sam: On liking music. *Cognition and Emotion* 1, 3 (1987), pp. 259–282.

25. Mark Grimshaw. 2008. Sound and immersion in the first-person shooter. *International Journal of Intelligent Games and Simulation* 5, 1 (2008), pp. 119–124.

26. Mark Grimshaw. 2009. The audio Uncanny Valley: Sound, fear and the horror game. In *Proceedings of the Audio Mostly Conference*. pp. 21–26.

27. Charlene Jennett, Anna L Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, 9 (2008), pp. 641–661.

28. 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 (CHI '16)*. ACM, New York, NY, USA, 3450–3461. DOI: http://dx.doi.org/10.1145/2858036.2858172

29. Alyx Jones. 2015. The Vanishing of Ethan Carter Game Audio Review. The Sound Architect, http://www.thesoundarchitect.co.uk/the-vanishing-of-ethan-carter-game-audio-review/. Accessed: 9th September 2017. (2015).

30. Kristine Jørgensen. 2008. Left in the dark: playing computer games with the sound turned off. *Collins, K. (ed.): From Pac-Man to Pop Music. Interactive Audio in Games and New Media. Ashgate, Farnham, United Kingdom* (2008), pp. 163–176.

31. Kristine Jørgensen. 2010. Time for new terminology? Diegetic and nondiegetic sounds in computer games revisited. *Grimshaw, Mark (ed.): Game Sound Technology and Player Interaction: Concepts and Developments. IGI global* (2010), pp. 78–97.

32. Philip Kollar. 2014. The Vanishing of Ethan Carter review: disappearing trick. Polygon, https://www.polygon.com/2014/10/10/6960363/the-vanishing-of-ethan-carter-review-pc. Accessed: 9th September 2017. (2014).

33. Martijn J.L. Kors, Gabriele Ferri, Erik D. van der Spek, Cas Ketel, and Ben A.M. Schouten. 2016. A Breathtaking Journey. On the Design of an Empathy-Arousing Mixed-Reality Game. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '16)*. ACM, New York, NY, USA, 91–104. DOI:http://dx.doi.org/10.1145/2967934.2968110

34. Daniel Kromand. 2008. Sound and the diegesis in survival-horror games. *Proceedings of the Audio Mostly Conference* (2008).

35. Pontus Larsson, Aleksander Väljamäe, Daniel Västfjäll, Ana Tajadura-Jiménez, and Mendel Kleiner. 2010. Auditory-induced presence in mixed reality environments and related technology. In *The Engineering of Mixed Reality Systems*. Springer, pp. 143–163.

36. Mats Liljedahl. 2011. Sound for fantasy and freedom. *Grimshaw, Mark (ed.): Game sound technology and player interaction: Concepts and developments. IGI global* (2011), pp. 22–44.

37. Jih-Hsuan Tammy Lin. 2017. Fear in virtual reality (VR): Fear elements, coping reactions, immediate and next-day fright responses toward a survival horror zombie virtual reality game. *Computers in Human Behavior* 72 (2017), pp. 350–361.

38. Scott D Lipscomb and Sean M Zehnder. 2004. Immersion in the virtual environment: The effect of a musical score on the video gaming experience. *Journal of Physiological Anthropology and Applied Human Science* 23, 6 (2004), pp. 337–343.

39. Matthew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication* 3, 2 (1997), pp. 1–43.

40. Teresa Lynch and Nicole Martins. 2015. Nothing to fear? An analysis of college students' fear experiences with video games. *Journal of Broadcasting & Electronic Media* 59, 2 (2015), pp. 298–317.

41. Zahira Merchant, Ernest T Goetz, Lauren Cifuentes, Wendy Keeney-Kennicutt, and Trina J Davis. 2014. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education. *Computers & Education* 70, C (2014), pp. 29–40.

42. David C Moffat and Katharina Kiegler. 2006. Investigating the effects of music on emotions in games. *Proceedings of the Audio Mostly Conference* (2006).

43. Lennart E. Nacke and Mark Grimshaw. 2011. Player-game interaction through affective sound. *Grimshaw, Mark (ed.): Game Sound, Technology and Player Interaction* (2011).

44. Lennart E. Nacke, Mark N. Grimshaw, and Craig A. Lindley. 2010. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers* 22, 5 (2010), pp. 336–343. DOI: http://dx.doi.org/10.1016/j.intcom.2010.04.005

45. Jim R Parker and John Heerema. 2008. Audio interaction in computer mediated games. *International Journal of Computer Games Technology* 2008 (2008), p.1.

46. Thomas D Parsons and Albert A Rizzo. 2008. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry* 39, 3 (2008), pp. 250–261.
47. James Robb, Tom Garner, Karen Collins, and Lennart E. Nacke. 2017. The Impact of Health-Related User Interface Sounds on Player Experience. *Simulation & Gaming* 48, 3 (2017), pp. 402–427. DOI: http://dx.doi.org/10.1177/1046878116688236
48. Thomas B Sheridan. 1992. Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments* 1, 1 (1992), pp. 120–126.
49. Chek Tien Tan, Tuck Wah Leong, Songjia Shen, Christopher Dubravs, and Chen Si. 2015. Exploring Gameplay Experiences on the Oculus Rift. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15)*. ACM, New York, NY, USA, 253–263. DOI: http://dx.doi.org/10.1145/2793107.2793117
50. Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments* 9, 5 (2000), pp. 497–503.
51. Jan Van Looy, Cédric Courtois, Melanie De Vocht, and Lieven De Marez. 2012. Player identification in online games: Validation of a scale for measuring identification in MMOGs. *Media Psychology* 15, 2 (2012), pp. 197–221.
52. David Watson, Lee A Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology* 54, 6 (1988), p.1063.
53. Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (1998), pp. 225–240.
54. Jiulin Zhang and Xiaoqing Fu. 2015. The Influence of Background Music of Video Games on Immersion. *Journal of Psychology & Psychotherapy* 5, 4 (2015), p.1.