Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS | ✓ Check for updates

The effect of audio on the experience in virtual reality: a scoping review

Isak de Villiers Bosman 🅳[a,b], Oğuz 'Oz' Buruk 🅳[b], Kristine Jørgensen 🅳[c] and Juho Hamari 🅳[b]

[a]Virtual Reality and Interaction Lab, Department of Information Science, University of Pretoria, Pretoria, South Africa; [b]Gamification Group, Faculty of Technology and Communication Sciences, Tampere University, Tampere, Finland; [c]Department of Information Science and Media Studies, University of Bergen, Bergen, Norway

ABSTRACT
The use of virtual reality (VR) has seen significant recent growth and presents opportunities for use in many domain areas. The use of head-mounted displays (HMDs) also presents unique opportunities for the implementation of audio feedback congruent with head and body movements, thus matching intuitive expectations. However, the use of audio in VR is still undervalued and there is a lack of consistency within audio-centedd research in VR. To address this shortcoming and present an overview of this area of research, we conducted a scoping review ($n = 121$) focusing on the use of audio in HMD-based VR and its effects on user/player experience. Results show a lack of standardisation for common measures such as pleasantness and emphasize the context-specific ability of audio to influence a variety of affective, cognitive, and motivational measures, but are mixed for presence and generally lacking for social experiences and descriptive research.

1. Introduction

The use of audio in virtual reality (VR) applications presents opportunities for novel uses and has received considerable attention (Kiss et al. 2020; Serafin et al. 2018). Audio can serve several functions in VR applications, such as providing action sounds based on interactions (Serafin et al. 2018) which can be sourced/localized to specific virtual objects (Kiss et al. 2020) and can be used to direct users' gaze and attention (Eames 2019; McArthur et al. 2017; Ulsamer et al. 2020). Audio can also create environment sounds or soundscapes that provide information about the surrounding virtual environment (Serafin et al. 2018) and play a crucial role in constructing perceived virtual spaces (Nordahl and Nilsson 2014). However, while the importance of audio for the overall VR user experience is recognized (Kern and Ellermeier 2020; Serafin et al. 2018), there is still a lack of secondary research which reviews the existing corpus of research regarding the effects of audio strategies on the user experience (UX) and player experience (PX) of VR applications.

In the broader context of human–computer interaction (HCI), audio has received considerable attention for enhancing a range of contexts through the sonification of interfaces and information (Papadopoulos et al. 2016; Sanchez 2010). There are many advantages of using audio as part of interaction mechanisms, such as the fact that audio feedback is often linked with shorter interaction times than visual feedback (Westerberg and Schoenau-Fog 2015), emits feedback omnidirectionally and constantly since we cannot close our ears (Serafin et al. 2018), does not clutter the visual real-estate (Palani and Giudice 2014), and can cater to users with visual impairments (Guillen et al. 2021; Rocchesso et al. 2019).

In the context of audio delivery and manipulation, VR affords unique interactions through catering for more intuitive interactions, i.e. sensorimotor contingencies (SCs) than typical desktop computer environments, such as using head tracking as an input mechanism for perception within a 3-D environment (Slater 2009). These SCs afford unique interactions relating to the psychoacoustic perception of audio, such as the fact that users of VR can use head movements to help localise audio in 3-D virtual spaces (Noisternig et al. 2003) which can lead to an enhanced sense of 'kinesonic congruence' between physical actions and resulting audio cues (Collins 2013, 57). Audio can also help solve some of the problems of content delivery in headset-based VR, such as the inability for designers to frame content visually due to freedom of head movement, which can be done with audio instead (McArthur et al. 2017).

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

While previous work in VR has investigated the effects of various qualities of visual displays on UX (Dickinson et al. 2019; Latoschik et al. 2019; Men et al. 2017), comparatively little focus has been given to the audio modality, which is often neglected in the development of various media projects (McArthur et al. 2017). There is still also a lack of clear design guidelines that consider the effects of various aspects of audio on UX in different contexts for VR applications (Kelling et al. 2018; Kiss et al. 2020; McArthur et al. 2017) and, as such, there is a general consensus that the audio modality is still under-utilised in VR (Serafin et al. 2018).

Previous review papers in the broad area of virtual environments and audio have investigated the role of audio in games (Guillen et al. 2021) as well an overview of audio games, i.e. games driven primarily by audio feedback (Urbanek and Güldenpfennig 2019). Specific to VR, review papers have investigated the UX of VR systems in a healthcare context (Mäkinen et al. 2022) and more generally (Kim, Rhiu, and Yun 2020). However, neither of the aforementioned papers focused on the role of audio in UX. Thus, this points toward a gap in comprehensive studies that analyze the field of VR in terms of audio effects on the user experience of VR. Filling this gap could benefit from pointing out the domain areas where audio is currently being studied which, in turn, provides guidance on areas where such a focus is lacking. Similarly, investigating which aspects of user/player experience are being studied and how the introduction of audio feedback affects these aspects would provide insight in a similar vein. Finally, such knowledge would provide implications for researchers and practitioners of audio-focused VR, such as pointing toward opportunities for more focused systematic reviews (Arksey and O'Malley 2005), as well as insight for practitioners as to how audio might be used in VR applications. Therefore, this article aims to provide a basis for the utilisation of sound in VR by using a scoping review to identify the range of contexts and outcomes being studied and how these are being studied (Munn et al. 2018) as well as an overview of the experiential outcomes

The literature review was guided by the following research questions:

(1) In which contexts are the effects of audio in HMD-based VR being studied?
(2) Which aspects of experience are being studied with regards to audio in HMD-based VR research?
(3) What insights does the literature research provide about the effect of audio on HMD-based VR experiences?

(4) What implications from the literature can be drawn regarding the design of future HMD-based VR audio research?

RQ 1–3 all pertain to the current state of the research and are addressed in the Results and Discussion sections, whereas RQ4 pertains to future work and research agenda regarding audio in VR and is answered in the Future Agenda.

1.1. Defining VR

Although interest in VR has grown significantly since its inception, there is still little consensus on what the term refers to (Bujic and Hamari 2020; Høeg et al. 2021). For example, the term is often used to refer to 'immersive' technologies that utilise head-mounted displays (HMDs) (Hruby 2019; Rogers et al. 2018) or CAVE-based systems (Sahai et al. 2012), but also to more traditional desktop-and-screen systems, e.g. (Buele et al. 2020; Lim et al. 2017), and some have even posited that a wide variety of media, such as film, photography, and music can also be considered a type of 'virtual reality' (Blascovich and Bailenson 2011, chap. 3; van Elferen 2011). Furthermore, it is also possible to mix reality and virtuality to varying degrees in different modalities, such as providing a visual augmented reality and an aural virtual reality (Larsson et al. 2010). Thus, to define our scope more clearly, we use the term VR to refer to systems that use a variety of display and tracking devices to allow for a set of intuitive interactions, i.e., valid SCs, thus facilitating a degree of natural interaction with a VE, and provides the user with feedback relevant to the VE in any number of modalities (Larsson et al. 2010; Nordahl and Nilsson 2014; Slater 2009). Furthermore, we limit our scope to that of VR created by HMDs for the reasons explained below.

1.2. Audio in VR

Perhaps the most frequently discussed topic in the context of VR audio is that of binaural or spatial audio, which refers to audio that is delivered in such a way as to mimic human perception of audio in real 3-dimensional environments, thus providing each 3-D object with a localisable 3-D audio source (Geronazzo et al., 2013). The use of binaural audio in VR presents opportunities for novel uses of audio, such as providing virtual objects with their own localisable audio sources, which can make better use of the spatiality of VR environments without being as invasive as visual feedback (Kiss et al. 2020). Spatialization of this kind is enabled by the use of high-fidelity tracking technology which

caters for valid SCs in the form of moving one's head and hearing the resulting change in audio (Noisternig et al. 2003). One audio-focused area of concern where this might be valuable is that of soundscape studies which focuses on the impact of sound and acoustics on the experiential properties of architectural and urban spaces (Brambilla and Di Gabriele 2009; Oberman, Bojanić Obad Šćitaroci, and Jambrošić 2018). This includes the investigation of different types of sounds, such as natural vs. artificial, and what effects they might have on human experiences, such as annoyance or restorativeness (Harriet and Murphy 2015; Uebel et al. 2021).

The use of VR in various contexts presents many opportunities but also new challenges to overcome, some of which can be addressed by the effective use of audio (McArthur et al. 2017; Serafin et al. 2018). In a narrative context where users/players are free to look around their environment, audio cues can be used to direct users' gaze and attention to critical points (Eames 2019; McArthur et al. 2017; Ulsamer et al. 2020), prevent cluttering the limited visual real-estate (Kiss et al. 2020; Shaw et al. 1993), and to provide a constant stream of information to be used for discerning spatial attributes of the environment itself (Nordahl and Nilsson 2014). Sound also possesses the property of being 'always turned on' (Larsson et al. 2010; Nordahl and Nilsson 2014), the ability to convey spatial characteristics, such as through reverberation (Larsson et al. 2010), and an omnidirectional quality (Jørgensen 2006; Turner et al. 2009) which does not rely on the limited visual field of view of display devices such as HMDs (Collins and Kapralos 2012). In contrast to visual feedback which might be static, the temporal nature of sound creates feedback which is always changing and 'in motion' (Larsson et al. 2010; Serafin et al. 2018). Combined with the spatial properties of sound, this is especially relevant for facilitating the 'place illusion', i.e., the illusion of finding oneself in a different physical environment, with minimal cognitive effort (Slater 2009; Turner et al., 2009). This temporal quality also means that our ability to judge temporal differences is much more precise in the audio than in the visual domain (Serafin 2020). However, while these place-like qualities of effective audio feedback provide the potential for sound to be a critical factor for users to feel present in VEs, the extent of the effect and of different audio types remains unclear (Kern and Ellermeier 2020).

Conversely, some of the properties of VR technology have the potential to be advantageous in various audio-driven contexts, such as music listening (Janer et al. 2016) and music composition (Turchet, Carraro, and Tomasetti 2022). For music listening, VR offers the potential for immersive experiences of recorded performances and could provide a degree of flexibility, such as focusing on specific instruments in an orchestra (Janer et al. 2016). For music and soundscape composition, the immersive aspect of VR has been found to increase the enjoyability and engagement of the experience (Turchet, Carraro, and Tomasetti 2022) and presents the possibility of utilising the technology's gestural tracking technology to extend the limited range of gesture-controlled instruments today (Cavdir and Wang 2022) as well as the flexibility to customise one's environment (Schlagowski et al. 2022). However, these areas are still relatively underexplored and lack clarity on how best to utilise VR (Schlagowski et al. 2022; Turchet, Carraro, and Tomasetti 2022). A logical starting point for addressing such concerns is a deeper understanding of the experience created by audio feedback in VR.

The tight coupling between interaction in the form of head movements and perceiving a change in the environment is principally tied to one of the most-discussed topics in VR research, i.e. that of presence (Kern and Ellermeier 2020; Nordahl and Nilsson 2014) since VR users adopt an 'egocentric virtual body representation' (Slater and Usoh 1993) or a sense of 'body ownership' (Nordahl and Nilsson 2014). This allows them to experience the virtual environment as if from a virtual body representation and allows for a sense of 'kinesonic congruity' between embodied player action and resulting action sounds (Collins 2013, p. 63). For this reason, the current study will focus on the use of HMDs with head-tracking, since their high-fidelity tracking allows for valid SCs in the form of head movements and more closely resembles a natural listening experience (Olko et al. 2017; Ooi et al. 2017) and generally focus on gestural input which may strengthen users' identification with virtual characters/avatars (Isbister 2016, 102). It should be noted that since the focus of this research is on the use of audio, we also include audio-only VR solutions provided they make use of head-tracking and adapt audio accordingly, e.g. (Garcia et al. 2015).

1.3. Shortcoming in VR audio research

Although there has been a significant growth in interest regarding audio in HCI (Rocchesso et al. 2019), in VR (Kiss et al. 2020; Serafin et al. 2018), and games (Collins 2020; van Elferen 2020), there is still much confusion and contradiction regarding how the use of audio affects user- and player experience while using HMD-based VR (Kapralos et al. 2017; Kern and Ellermeier

2020; Rogers 2017; Rogers et al. 2020). There are many possible reasons for this, including the novelty bias associated with newer technologies such as HMDs (Rogers et al. 2020) and limited understanding of various concepts in audio design (Jørgensen 2011; Ribeiro et al. 2020), especially pertaining to VR (Jain et al. 2021).

Research in VR audio is often limited to pure performance-driven metrics, such as being able to accurately localise the position of audio in 3-D space, at the expense of other perceptual/experiential measures that might benefit from further investigation, both quantitative and qualitative (Jenny and Reuter 2021). There is also a general lack of comparability between different studies with regards to differences in research design, such as how audio cues were presented, which aspects were studied, what technologies were used, and the differing roles played by audio (Jørgensen 2017; Kern and Ellermeier 2018; Kern and Ellermeier 2020; Kern et al. 2020). Audio classification also tends to focus on the characteristics of sound sources at the expense of considering other factors that could be useful, such as the intent (Jain et al. 2021). By investigating the use of audio in various HMD-based VR contexts and conducting secondary research on aspects concerning the experiences created by these contexts, we aim to provide insight into the aspects of user- and player experience being studied and how this has been done, as well to provide a basis for future research in the field of audio for HMD-based VR.

2. Method

This study uses a scoping review for assessing the effect of HMD-based audio on user experience in VR. The goal of the review is to present and analyze the current state of the art in this area using a protocol that allows for rigour and accountability. More specifically, a scoping review aims to map the findings from an often-heterogeneous area of research to provide an overview of the range of research and to point out gaps (Arksey and O'Malley 2005). A scoping review was thus selected as the broader topic of audio-driven VR has not been explored in previous reviews and, as such, a broader approach which provides an entry point into this area and takes on an exploratory approach was deemed appropriate (Chang 2018). The following three sections detail the protocol with regards to search and screening of articles and analysis and presentation of data, which was based on the PRISMA-ScR checklist for scoping reviews (Tricco et al. 2018).

2.1. Search strategy

A search string was created which targeted the three main areas being investigated, namely VR, audio, and user/player experience. To cover the various ways audio feedback is referred to in the literature, a variety of audio-related terminology was used in the search string, as shown below. Furthermore, in order to include the widest possible interpretation of the concepts of 'user experience' and 'player experience', the more general term 'experience' was used instead. The two databases targeted were Scopus and ACM DL and the search strings retrieved matches from the title, abstract, and/or keywords. The exact search strings used for each database were as follows:

Scopus:

TITLE-ABS-KEY ((VR OR (virtual AND reality)) AND (sound* OR music* OR audio* OR audito* OR aural* OR acoustic* OR sonic*) AND (experience* OR UX OR PX)) AND (LIMIT-TO (DOCTYPE,"cp") OR LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"ch"))

ACM:

(Abstract: (VR OR (virtual AND Reality)) OR Title: (VR OR (virtual AND Reality)) OR Keywords: (VR OR (virtual AND Reality)))

AND

(Abstract: (sound* OR music* OR audio* OR audito* OR aural* OR acoustic* OR sonic*) OR Title: (sound* OR music* OR audio* OR audito* OR aural* OR acoustic* OR sonic*) OR Keywords: (sound* OR music* OR audio* OR audito* OR aural* OR acoustic* OR sonic*))

AND

(Abstract: (experience* OR UX OR PX) OR Title: (experience* OR UX OR PX) OR Keywords: (experience* OR UX OR PX))

The ACM search string was formulated in such a way as to return similar results as the Scopus string. Since the ACM DL advanced search does not allow searching for title, abstract, and keyword using a single filter, the above string was used to achieve the same result. For items retrieved from ACM DL, document types were checked during the manual exclusion phase. The searches were conducted on 08 September 2020.

2.2. Review procedure

The following were selected as selection criteria on which to include or reject retrieved papers:

- The paper is published in a peer-reviewed venue
- The paper is in English
- The full paper is available
- The paper discusses the influence of audio on user or player experience in the context of HMD-based virtual reality. Checking whether papers are in scope thus involved three criteria:
  - (a) VR: the study made use of or discussed the use of HMDs or, for purely audio-based solutions, head-tracking systems
  - (b) Audio: the effects of audio were studied on their own, either quantitatively or qualitatively
  - (c) Experience: any aspect of user or player experience were studied empirically

The procedure for finalising the list of articles to analyze was as follows: first, the search string was used to retrieve all records from the aforementioned databases, which were compiled into a single file and duplicates were removed, leaving 1830 papers. This list was then screened using the title and abstract of each publication using the selection criteria listed above, which resulted in 987 articles being excluded. Of these rejections, 25 were for records that were identified as not being research papers (criteria 1), such as calls for proposals, while the rest were for not fitting the scope of the review (criteria 4).

The remaining list of publications was then assessed using the full text, again based on the selection criteria. In the case where we did not have access to the publication, we contacted the authors via email or Research-Gate; 33 publications were accessed this way. During the full-text assessment, an additional 722 publications were excluded, 21 being for records that were not full research papers, such as extended abstracts (criteria 1), 13 based on the publications not being in English (criteria 2), 45 based on us not being able to access these

Table 1. Extracted data from articles.

Data	Description
Publication venue	Type of venue (journal article, conference paper, etc.) and name of the publication (title of journal, conference, etc.)
Domain area	The domain area of study in which the study was performed (if any)
Study design	Design(s) employed for study, e.g. randomised controlled trial
Data collection tools	Tools/methods used for data collection
Analysis methods	Statistical methods for quantitative data and qualitative methods for qualitative data
Measured outcomes + questionnaires/scales	Experiential outcomes which were measured, and questionnaires/scales used to measure these
Effect of audio	Reported outcomes of study pertaining to experiential outcomes

publications (criteria 3), and the remaining 643 were for not fitting the scope of the review (criteria 4). This left 121 publications to be used for analysis.

2.3. Analysis

The list of variables extracted from each article is described in Table 1 as per the PRISMA-Scr requirements (Tricco et al., 2018):

These variables were selected to provide a broad overview of the area of study and to be able to point out possible gaps/shortcomings. The publication venue and domain area indicate the context in which each was conducted, which was expected to provide insights into areas where audio-driven research might be lacking. Study design, data collection tools, and analysis methods provide an overview of methodological concerns that guide the research and can provide specific suggestions for addressing such concerns in future research. Measured outcomes, questionnaires/scales, and effect of audio specify the topical focus of the various studies and the insights gleaned from each which point toward experiential outcomes where audio is effective, ineffective, unexplored, etc.

The variables were extracted from each article by the first author. During the process, there was frequent communication between authors to discuss and resolve issues relating to scope, data coding, and reporting. The synthesis of results involved an initial open coding process for all variables, except publication venue and year, and categories were inductively created for each variable through author discussion in a similar process as recent review papers, e.g. (Pater et al. 2021). This was followed by axial coding (Scott and Medaugh 2017) for the domain areas, measured outcomes, and effects of audio where codes were grouped into higher-level categories and existing classifications were retroactively updated. Finally, the data were compiled into tables using these categories, which were used to derive insights and implications.

3. Results

3.1. Publication year

As can be seen from Figure 1 below, the number of studies that focus on the use of audio in VR applications has seen a significant growth over the past 20 years. It must also be noted that the database searches were conducted in September 2020 which significantly lowers the number of citations for the year 2020 Figure 2.
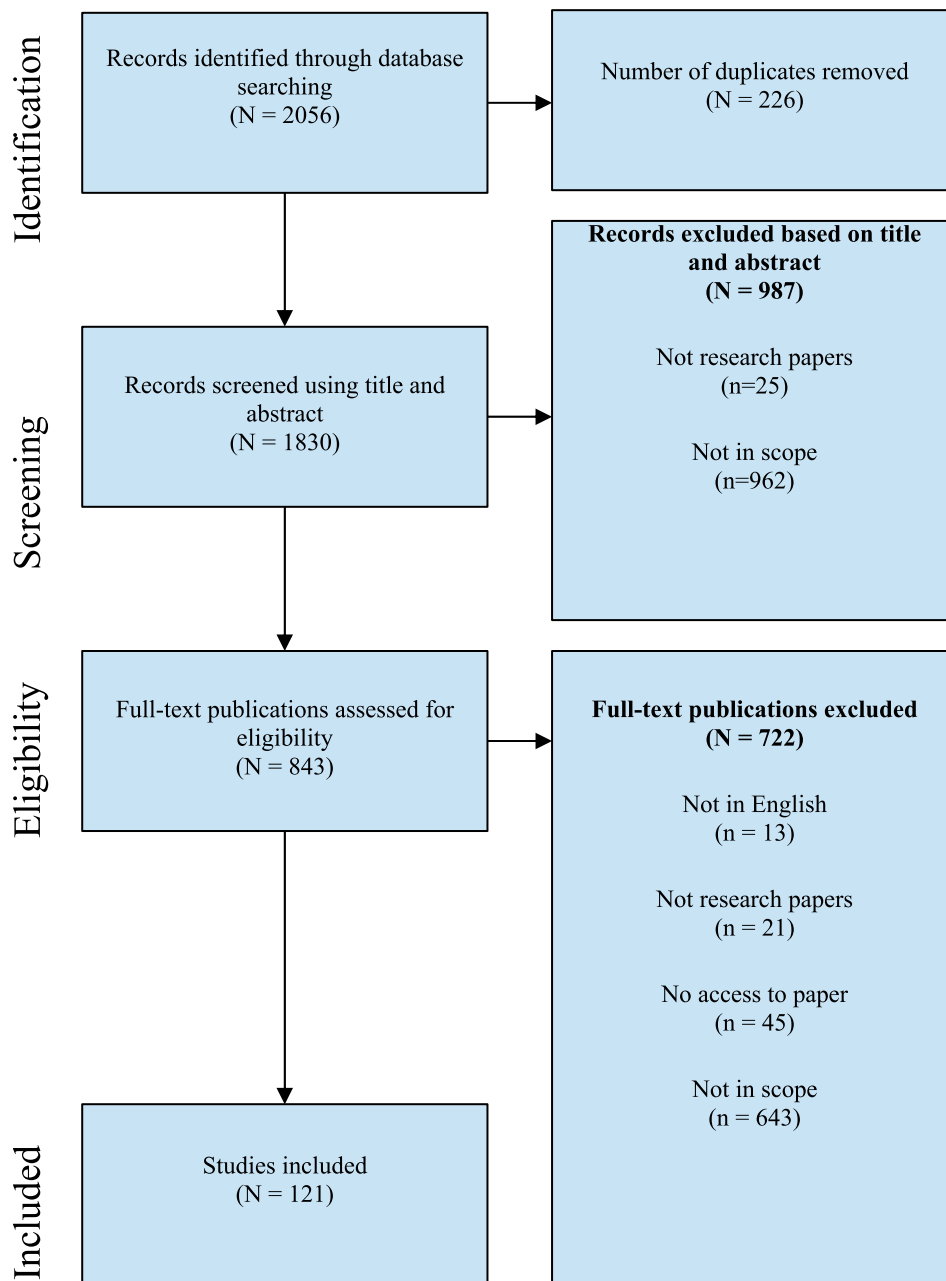
Figure 1. Process diagram of retrieval and eligibility screening of papers.

3.2. Publication venue

The breakdown of publication venues for the final list of included publications is listed in Table 2.

The corpus was drawn from 83 unique venues. The most frequent were the INTER-NOISE conference (8 papers), ACM CHI conference (5 papers), IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (5 papers), Liebert Cyberpsychology, Behavior, and Social Networking journal (4 papers), IOS Press Studies in Health Technology and Informatics (3 papers), IEEE VR Workshop on Sonic Interactions for Virtual Environments (SIVE) (3 papers), and Elsevier

Landscape and Urban Planning journal (3 papers). All other venues had either two or one paper per venue. The corpus did not include any book chapters, possibly due to the inclusion requirement pertaining to empirical research of one/more experiential aspects.

3.3. Domain areas

The list of domain areas for all included publications is shown in Table 3. Note that the list of domain areas was created inductively and iteratively during the coding process through discussion between all authors and
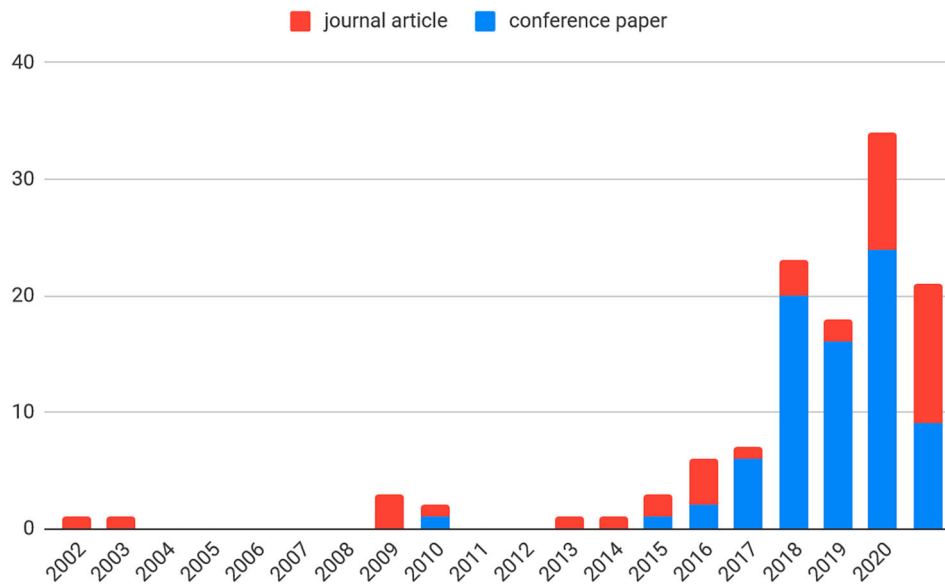
Figure 2. Number of publications for each publication venue per year.

that some papers fall into more than one domain category.

The most frequent area for studying the effects of audio on VR experiences is for the purposes of entertainment. Likely due to the focus on audio, this is most often done in the realm of studies related to music, such as enabling new techniques or interfaces for musical expression (Chang, Kim, and Kim 2018; Deacon, Stockman, and Barthet 2017) as well as for music education in the form of ear training (Fletcher, Hulusic, and Amelidis 2019; Pedersen et al. 2020), and enabling new ways of viewing and listening to recorded music (Droste, Letellier, and Sieck 2018; Holm, Väänänen, and Battah 2020; Kasuya et al. 2019; Lind et al. 2017; Rees-Jones and Daffern 2019; Shin et al. 2019). The corpus also contains a relatively small number of studies that focus specifically on audio and player experiences within VR games (Bombeke et al. 2018; Rogers et al. 2018; Wedoff et al. 2019), although it is possible that this number stems from the choices of databases consulted. There is also a focus on audio in other types of narrative experiences, such as experiences that are linear and video-based (Dumlu and Demir 2020; Gospodarek et al. 2019; Vosmeer and Schouten 2017) or interactive and audio-driven (Geronazzo et al. 2019), and finally one instance using novel

interaction techniques to create an immersive scuba diving experience (Jain et al. 2016).

Another popular area for studying the effects of audio on VR experiences is for the purpose of soundscape studies. Soundscape research in this context often generally focuses on the effects of different types of sounds in architectural and urban contexts, using VR as a tool for visually and aurally recreating various environments. This includes focusing on the effects of artificial noise in urban environments (Aalmoes and Zon 2019; Cranmer et al. 2020; Echevarria Sanchez et al., 2017a, 2017b; Guang-yu, Yu-an, and Haochen 2019; Jeon and Jo 2019; Letwory, Aalmoes, and Miltenburg 2018; Maffei et al. 2013; Masullo, Maffei, and Pellegrino 2019; White et al. 2014) and, to a much lesser extent, focusing on natural soundscapes (Lindquist et al. 2020).

Where audio is studied in the broader domain of psychotherapy, the most frequently studied use for VR is for exposure therapy, such as for the treatment of post-traumatic stress disorder (PTSD) (Difede and Hoffman 2002; Josman et al. 2008), phobias (Brinkman et al. 2009; North et al. 2008), and disorders related to sensory processing (Johnston, Egermann, and Kearney 2020; Rossi et al. 2018), as well as the general study of underlying processes related to exposure therapy (Andreano et al. 2009). Mindfulness practice is also a topic of key interest, such as in a rehabilitative or analgesic context (Gomez et al. 2017; Nararro-Haro et al. 2016; Pizzoli et al. 2019) or more generally in investigating how VR technologies might contribute to mindfulness practice (Prpa et al. 2017, 2018; Seabrook et al. 2020). The experiential role of audio in VR is also studied in various therapeutic and/or rehabilitative

Table 2. Publication venues of included papers.

	n	%
Conference paper	79	65.29
Journal article	42	34.71
Total	121	100

Table 3. Domain areas of included papers.

Domain	References	n	%
Entertainment	Bombeke et al. (2018); Chang, Kim, and Kim (2018); Deacon, Stockman, and Barthet (2017); Droste, Letellier, and Sieck (2018); Dumlu and Demir, (2020); Fletcher, Hulusic, and Amelidis (2019); Geronazzo et al., (2019); Gospodarek et al., (2019); Holm, Väänänen, and Battah (2020); Jain et al., (2016); Kasuya et al. (2019); Lind et al. (2017); Pedersen et al. (2020); Rees-Jones and Daffern (2019); Rogers et al. (2018); Shin et al. (2019); Vosmeer and Schouten (2017); Wedoff et al. (2019)	19	15.2
Soundscape studies	Aalmoes and Zon (2019); Cranman et al. (2020); Echevarria Sanchez et al. (2017a, 2017b); Guang-yu, Yu-an, and Haochen (2019); Hong et al. (2019); Jeon and Jo (2019); Jo and Jeon (2020); Letwory, Aalmoes, and Miltenburg (2018); Li and Xie, (2018); Lindquist et al., (2020); Maffei et al. (2013); Masullo, Maffei, and Pellegrino (2019); Ong et al., (2017); Ooi et al., (2017); Park et al. (2020); White et al. (2014)	18	14.4
Psychotherapy/ rehabilitation	Andreano et al., (2009); Brinkman et al., (2009); Byrns et al. (2020); Difede and Hoffman, (2002); Gomez et al. (2017); Johnston, Egermann, and Kearney (2020); Josman et al. (2008); Nararro-Haro et al. (2016); North et al. (2008); Optale et al., (2001); Park et al., (2020); Pizzoli et al. (2019); Prpa et al. (2017, 2018); Rossi et al., (2018); Seabrook et al. (2020); Sorkin et al., (2008); Zeroth, Dahlquist, and Foxen-Craft (2019)	18	14.4
Arts and cultural heritage	Bialkova and Gisbergen (2017); Falconer et al., (2020); Kelling et al. (2018); Y. Li, Tennent, and Cobb (2019); Prpa et al. (2017, 2018); Rzayev et al. (2019); Sprung et al. (2018); Tennent et al. (2020)	9	7.2
Physical therapy/exercise	Baka et al., (2018); Bird et al. (2019, 2020)	3	2.4
Engineering	Barlow et al., (2019); Michailidis et al., (2019); Yin et al., (2019)	3	2.4
Gamification	Hicks et al., (2019); Steadman et al., (2019)	2	1.6
Built environment	Ishikawa, (2019); Wu et al. (2018)	2	1.6
Cognitive psychology		1	0.8
Human-robot interaction	Li, Tennent, and Cobb (2019)	1	0.8
Culinary	Ibrahim and Wan Bashir, (2019)	1	0.8
Accessible/assistive technology[a]	Bălan et al. (2018); Dong, Wang, and Guo (2017); Picinali et al. (2014); Siu et al. (2020); Souza Silva et al. (2019); Wedoff et al. (2019)	6	4.8
Audio in VR[a]	Brinkman, Hoekstra, and van Egmond (2015); Davies et al., (2017); Geronazzo et al., (2018); Hong et al., (2018); Huang et al., (2019); Kurabayashi et al., (2014); McArthur (2016); Mehra et al., (2015); Moraes et al. (2020); Olko et al. (2017); Pelegrin Garcia et al., (2015); Rummukainen et al., (2017); Schoeffler et al. (2015); Steadman et al., (2019); Stecker et al., (2018); Suarez et al. (2017); Ulsamer et al., (2020); Yan, Wang, and Li (2019)	18	14.4
No domain[b]	Çamci (2019); Chen et al. (2017a, 2017b, 2018); Chirico and Gaggioli (2019); Chittaro (2012); Feng, Dey, and Lindeman (2016); Gao, Kim, and Kim (2018); Ghosh et al. (2018); Kruijff et al. (2016); Lee, Bruder, and Welch (2017); Lee and Lee (2017); Liao et al. (2020); Narciso et al. (2019); Oh, Herrera, and Bailenson (2020); O'Hagan, Williamson, and Khamis (2020); Peng et al. (2020); Sawada et al. (2020); Shimamura et al. (2020); Sikström, de Götzen, and Serafin (2015, 2016a, 2016b); Van den Broeck, Pamplona, and Fernandez Langa (2017); Zhang et al. (2018); Zhao et al. (2017)	24	19.2
		125	100

[a]Does not discuss a specific domain per se, but focuses on specific topics within the general area of VR.
[b]Also does not discuss a specific domain, but rather VR research in general.

contexts (Byrns et al. 2020; Optale et al. 2001; Park et al. 2020; Sorkin et al. 2008; Zeroth, Dahlquist, and Foxen-Craft 2019). The enabling of physical activity and/or rehabilitation is also studied, albeit to a much lesser extent, specifically for physical rehabilitation through music therapy (Baka et al. 2018) and the effect of VR and music on the experience of exercise (Bird et al. 2019, 2020).

In the arts and cultural heritage domain, audio is used in the design of experiences that aim to emulate traditional arts/museum experiences (Bialkova and Gisbergen 2017; Kelling et al. 2018; Y. Li, Tennent, and Cobb 2019; Rzayev et al. 2019; Sprung et al. 2018; Tennent et al. 2020) as well as to create new immersive experiences such as exploring archeological sites (Falconer et al. 2020) and using breath-input for novel purposes (Prpa et al. 2017, 2018). In the domain of engineering, audio is used for simulating diagnostics (Barlow et al. 2019; Michailidis et al. 2019) and enabling multimodal feedback for assembly tasks (Yin et al. 2019). In a built environment context, VR is used to study building acoustics (Ishikawa 2019) and traffic crossings (Wu et al. 2018). There is also one instance of using VR to investigate experience in a culinary context (Ibrahim and Wan Bashir 2019).

Lastly, most papers that include some investigation of audio effects on the experience of VR did so in a broader context not specific to a single domain/discipline but rather as part of research into the medium of VR itself. This includes studies that focus on accessible or assistive technologies, including audio-only VR for people with visual impairments (Bălan et al. 2018; Picinali et al. 2014; Siu et al. 2020; Wedoff et al. 2019) and the investigation of age (Souza Silva et al. 2019) and visual impairments (Dong, Wang, and Guo 2017) with regards to audio feedback in VR. Many of these medium-specific studies have a focus on technical aspects of audio, investigating the experiential effects of technological factors such as 3D/binaural audio (Brinkman, Hoekstra, and van Egmond 2015; Davies et al. 2017; Hong et al. 2018; Huang et al. 2019; Olko et al. 2017; Rummukainen

et al. 2017; Schoeffler et al. 2015; Ulsamer et al. 2020; Yan, Wang, and Li 2019) and specific technological factors within this area such as HRTFs (Geronazzo et al. 2018; Suarez et al. 2017), head-tracking (Kurabayashi et al. 2014; Steadman et al. 2019), room acoustical properties (Garcia et al. 2015; Stecker et al. 2018), sound source displacement (McArthur 2016; Moraes et al. 2020), and different sound propagation techniques (Mehra et al. 2015). The 'no domain' remainder of these studies exploring the medium of VR refers to the large group of papers studying some other aspect of the VR medium.

3.4. Study design

The number of papers using different study designs is found in Table 4. Note that the total number of studies exceeds 121 since three papers respectively consisted of two (Rogers et al. 2018), two (Wu et al. 2018), and three (Ghosh et al. 2018) separate studies that were within the scope of this review and one paper (Shin et al. 2019) utilised a design combining a quasi-experiment and a randomised crossover.

The overwhelming majority of this research involves exposing participants to some type of technological intervention, either developed/recorded for the purposes of the study ($n = 110$) or using existing VR content ($n = 13$), and then, in the case of experimental or crossover studies, collecting quantitative data regarding the experience. Note that some studies utilised more than one VR application. The phenomenological studies listed also involved exposing participants to some type of technological intervention/exhibition but focused on qualitative experiential data. Descriptive studies investigated populations interacting with existing technologies, in other words, the studies did not involve exposing participants to any type of intervention. Finally, one study (Ghosh et al. 2018) employed a co-creative workshop design to collect qualitative data.

3.5. Data type

The list of data collection tools is given in Table 5.

Due to the focus on experimental and crossover studies, the majority of studies focused on collecting quantitative data and many studies used a combination of several questionnaires in a single study. In total there were 223 instances of using questionnaires used across all studies with 79 of them being existing questionnaires and the remaining 144 being custom created for the purposes of the study. Note that this number does not account for reuse of existing questionnaires, but it does assume that each custom questionnaire created to measure a single experiential outcome is unique. As such, a common approach was to create new quantitative self-report instruments, possibly due to the large number of experiential measures and the fact that many of these measures are not commonly studied.

Usage data broadly refer to data gathered from system use whereas performance data specifically refer to cases where usage was framed as having a better/worse outcome; both were only considered if their use was explicitly linked to some experiential outcome. Interviews and qualitative surveys were the most common instruments used for collecting qualitative data although observations were also commonly used in the form of direct observation or post-hoc analysis of audio and/or video material. The only use of focus groups was in (Ghosh et al. 2018) which also involved the generation and analysis of visual sketches as participant output from a co-creative workshop exercise.

3.6. Analysis methods

The list of analysis techniques for both quantitative and qualitative data as discussed in the corpus are respectively listed in Tables 6 and 7 below.

For quantitative data, the most common analysis methods relied on descriptive statistics and quantitative comparisons, owing to the large number of studies that developed multiple VR-based implementations and compared some experiential aspects between them.

Table 4. Study designs of included papers.

Type of study		N		%
Experimental	Randomised controlled trial	39	19	15.08
	Quasi-experiment		20	15.87
Crossover	Randomised crossover	63	24	19.05
	Counterbalanced crossover		18	14.29
	Quasi-crossover		21	16.67
Phenomenological		14		11.11
Case study		6		4.76
Descriptive		3		2.38
Workshop		1		0.79
Total		126		100

Table 5. Data collection tools used in the included papers.

Quantitative data			Qualitative data			Total
	n	%		n	%	
Quantitative survey	83	41.92	Interview	28	14.14	
Performance data	21	10.61	Qualitative survey	20	10.1	
Physiological data	16	8.08	Observation	14	7.07	
Usage data	14	7.07	Focus group	1	0.51	
			Sketches	1	0.51	
Total	134	67.68		64	32.33	198

Table 6. Quantitative analysis techniques used.

Quantitative analysis	n	%
Descriptive statistics	78	32.5
Quantitative Comparison	127	52.9
(T-tests, T-contrasts, paired sample t-test, ANOVA, ANCOVA, MANOVA, Kruskal-Wallis H-test, post-hoc pairwise Conover, pairwise comparison test, Tukey's HSD, Fisher exact test, Friedman test, Mann-Whitney U test, Wilcoxon signed-rank test)		
Quantitative Modeling	12	5
(Probit regression analysis, linear regression analysis, 2D bi-dimensional regression analysis, multinomial logistic regression, structural equation modelling, linear mixed modeling, repeated-measures mixed models, cumulative link model)		
Quantitative Association-based	20	8.3
(correlation, Pearson's correlation, non-parametric Pearson's correlation, Pearson's chi-squared test, Spearman rank correlation, chi-square test)		
Other quantitative statistics	3	1.3
(independent component analysis, Harman's one-factor test, Markov chain analysis)		
Total	240	100

Table 7. Qualitative analysis techniques used.

Qualitative analysis	N	%
Thematic Analysis	17	70.8
Grounded Theory	2	8.3
Linguistic Inquiry Word Count	1	4.2
Verbal Protocol Analysis	1	4.2
Interaction Analysis	1	4.2
Cinemetrics	1	4.2
Diagramming	1	4.2
Total	24	100

While eight studies used only descriptive statistics to infer results (Byrns et al. 2020; Falconer et al. 2020; Hong et al., 2018; Kelling et al. 2018; Michailidis et al., 2019; Yan, Wang, and Li 2019; Zhang et al. 2018; Zhao et al. 2017), the majority of studies used descriptive statistics in conjunction with other methods.

Analysis of qualitative data largely made use of thematic analysis or grounded theory to create common themes in data, with two studies using Cohen's kappa (Oh, Herrera, and Bailenson 2019; Wedoff et al. 2019) and one using Krippendorf's alpha (Jain et al. 2016) for reliability measures. Other approaches for verbal data included using a linguistic inquiry word count to measure affective valence (Oh, Herrera, and Bailenson 2019) and the use of verbal protocol analysis to generate semantic scales for audio attributes (Olko et al. 2017). Finally, three studies used observation data by analyzing these data based on user interactions (Deacon, Stockman, and Barthet 2017), based on the 'camera' movement generated by users (Dumlu and Demir 2020) and by creating and analyzing transition diagrams representing decisions made (Brinkman et al. 2009).

3.7. Measured outcomes

The experiential outcomes measured in the corpus are listed in Table 8 below. During the coding process, categories were inductively and iteratively derived from the outcomes listed in the corpus through discussion between the authors. Similarly, outcome variables that were considered similar, such as fear and anxiety, as well as being inversely equivalent, such as comfort and discomfort, were grouped inductively based on discussions between authors. Since this list specifically refers to discretely measurable outcomes, the list does not include qualitative data from the corpus. Where objective measures were explicitly linked to experience-outcomes, said outcome is listed.

When combining self-reported and other measures, the experiential outcome that was measured the most across all studies was that of presence ($n = 25$), followed by realism ($n = 15$), and immersion ($n = 13$). Pleasantness was also frequently measured ($n = 12$), which is largely attributable to the frequency of soundscape research in the corpus, as nine of the studies that include pleasantness as an outcome fall under this domain. While activation/arousal ($n = 10$) is measured with self-reports it is also the outcome that was most often measured using physiological measures, such as skin conductance level, galvanic skin response, electrodermal activity, etc.

For measures relating to emotion/affect, significantly more emphasis was placed on studying outcomes generally associated with positive affect than with negative affect. Of all instances of measuring emotion/affect, 71 outcomes were for positive affective measures, such as pleasantness, enjoyment, liking/affective attraction, etc. while 43 outcomes related more with negative outcomes, such as simulator sickness, fear/anxiety, tension/annoyance, etc. (with two outcomes more broadly investigating valence). Three common focus areas related to affect were that of perceived stimulation from audio stimuli, i.e. activation, arousal, eventfulness, excitement, and stimulation, that of perceived effects from interest in the stimuli, i.e. interest, novelty, and fascination, and that of calming effects of audio stimuli, i.e. relaxation and calmness.

Many perceptual measures where audio was studied were related to quality measures and how experiences conformed to expectations. This was commonly done using perceived realism or naturalness and one instance of believability. More general measures of perceived overall quality and preference might be said to relate more closely to personal preferences. Other measures in this category refer either to perceived aspects of the environment or of embodied experience. The former includes the effect of audio on perceived audio/visual

Table 8. Measured outcomes organised by categories related to experience.

Subjective self-report			Objective (performance/usage/physiological/linguistic)		
Outcome	References	n	Outcome	References	n
Emotion/affect (n = 114)					
Pleasantness	Echevarria Sanchez et al. (2017a, 2017b); Holm, Väänänen, and Battah (2020); Hong et al. (2019); Jo and Jeon (2020); Kelling et al. (2018); Masullo, Maffei, and Pellegrino (2019); Ong et al., (2017); Ooi et al., (2017); Park et al. (2020); Pizzoli et al. (2019)	12	Activation/arousal	Chen et al. (2017a); Chittaro (2012); McArthur (2016); Moraes et al. (2020); Park et al. (2020)	5
Positive affect	Bird et al. (2019, 2020); Chen et al. (2017b); Chirico and Gaggioli (2019); Fletcher, Hulusic, and Amelidis (2019); Geronazzo et al., (2019); Gomez et al. (2017); Nararro-Haro et al. (2016); Pedersen et al. (2020); Pizzoli et al. (2019); Rogers et al. (2018)	11	Fear/anxiety	Brinkman, Hoekstra, and van Egmond (2015); Johnston, Egermann, and Kearney (2020); Li and Xie, (2018)	3
Negative affect	Chen et al. (2017b); Chirico and Gaggioli (2019); Fletcher, Hulusic, and Amelidis (2019); Geronazzo et al., (2019); Gomez et al. (2017); Nararro-Haro et al. (2016); Pedersen et al. (2020); Pizzoli et al. (2019); Rogers et al. (2018)	9	Affective valence	Moraes et al. (2020); Oh, Herrera, and Bailenson (2019)	2
Fear/anxiety	Brinkman, Hoekstra, and van Egmond (2015); Chittaro (2012); Johnston, Egermann, and Kearney (2020); Lee, Bruder, and Welch (2017); North et al. (2008); Prpa et al. (2017); Rzayev et al. (2019); E. Sikström et al., (2016)	8	Discomfort	Aalmoes and Zon (2019)	1
Simulator sickness	Feng, Dey, and Lindeman (2016); Kruijff et al. (2016); Y. Li, Tennent, and Cobb (2019); Liao et al. (2020); Narciso et al. (2019); Peng et al. (2020); Sawada et al. (2020)	7	Frustration	Byrns et al. (2020)	1
Tension/annoyance	Fletcher, Hulusic, and Amelidis (2019); Guang-yu, Yu-an, and Haochen (2019); Jeon and Jo (2019); Maffei et al. (2013); Masullo, Maffei, and Pellegrino (2019); Pedersen et al. (2020); White et al. (2014)	7	Pain tolerance	Zeroth, Dahlquist, and Foxen-Craft (2019)	1
Liking/affective attraction	Bialkova and Gisbergen (2017); Lee, Bruder, and Welch (2017); Y. Li, Tennent, and Cobb (2019); Michailidis et al., (2019); Rzayev et al. (2019); Schoeffler et al. (2015)	6	Positive affect	Park et al. (2020)	1
Activation/arousal	Bird et al. (2019, 2020); Chen et al. (2017b); Pizzoli et al. (2019); Rogers et al. (2018)	5			
Discomfort	Brinkman, Hoekstra, and van Egmond (2015); Jo and Jeon (2020); Josman et al. (2008); North et al. (2008); Peng et al. (2020)	5			
Enjoyment	Bird et al. (2019, 2020); Gospodarek et al., (2019); Shin et al. (2019); Erik Sikström et al., (2016)	5			
Calmness	Jo and Jeon (2020); Ooi et al., (2017); Park et al. (2020)	4			
Dominance	Chen et al. (2017b); Pizzoli et al. (2019); Rogers et al. (2018)	3			
Eventfulness	Jo and Jeon (2020); Ooi et al., (2017)	3			
Excitement	Jo and Jeon (2020)	2			
Exertion	Bird et al. (2019, 2020)	2			
Novelty	Y. Li, Tennent, and Cobb (2019); Michailidis et al., (2019)	2			
Satisfaction	Oh, Herrera, and Bailenson (2019); Shimamura et al. (2020)	2			
Stimulation	Y. Li, Tennent, and Cobb (2019); Michailidis et al., (2019)	2			
Awe	Chirico and Gaggioli (2019)	1			
Distress	Difede and Hoffman, (2002)	1			
Fascination	Park et al. (2020), p.	1			
Interest	Kelling et al. (2018)	1			
Pain intensity	Zeroth, Dahlquist, and Foxen-Craft (2019)	1			
Relaxation	Pizzoli et al. (2019)	1			
Perceptual/embodiment (n = 48)					
Realism	Chang, Kim, and Kim (2018); Davies et al., (2017); Feng, Dey, and Lindeman (2016); Huang et al., (2019); Jeon and Jo (2019); Lindquist et al., (2020); Ooi et al., (2017); Peng et al. (2020); Rzayev et al. (2019); Sawada et al. (2020); E. Sikström et al., (2016); Yan, Wang, and Li (2019)	12	Realism	Barlow et al., (2019); Siu et al. (2020); Sorkin et al., (2008)	3
Naturalness	Bialkova and Gisbergen (2017); Chang, Kim, and Kim (2018); Shimamura et al. (2020); E. Sikström et al., (2016)	4	Spatial perception	Feng, Dey, and Lindeman (2016); Ulsamer et al., (2020)	2
Quality	Yan, Wang, and Li (2019); Zhang et al. (2018); Zhao et al. (2017)	3	Parasympathetic activity	Bird et al. (2020)	1

(*Continued*)

Table 8. Continued.

Subjective self-report			Objective (performance/usage/physiological/linguistic)		
Outcome	References	n	Outcome	References	n
Spatial perception	Brinkman, Hoekstra, and van Egmond (2015); Gospodarek et al., (2019); Kasuya et al. (2019)	3	Preference	Çamci (2019)	1
Synchrony of audio/visual stimuli	Chang, Kim, and Kim (2018); Geronazzo et al., (2018); Shimamura et al. (2020)	3			
Ownership of virtual body	Erik Sikström et al., (2016); Sikström, de Götzen, and Serafin (2015)	2			
Preference	Chen et al. (2017b); Park et al. (2020)	2			
Sense of movement	Feng, Dey, and Lindeman (2016); Kruijff et al. (2016)	2			
Audio/visual stimulus removal	Shimamura et al. (2020)	1			
Avatar weight	E. Sikström et al., (2016)	1			
Awareness of body movement	Erik Sikström et al. (2016)	1			
Believability	Gao, Kim, and Kim (2018)	1			
Biodiversity	Lindquist et al., (2020)	1			
Crowd density	Jo and Jeon (2020)	1			
Haptic feedback	Chang, Kim, and Kim (2018)	1			
Heart rate awareness	Chen et al. (2017b)	1			
Integration of visual elements	Masullo, Maffei, and Pellegrino (2019)	1			
Player identification	Rogers et al. (2018)	1			
Cognitive (n = 58)					
Presence	Bialkova and Gisbergen (2017); Brinkman, Hoekstra, and van Egmond (2015); Chirico and Gaggioli (2019); Deacon, Stockman, and Barthet (2017); Dong, Wang, and Guo (2017); Feng, Dey, and Lindeman (2016); Gao, Kim, and Kim (2018); Geronazzo et al., (2018); Holm, Väänänen, and Battah (2020); Jain et al., (2016); Lee, Bruder, and Welch (2017); Liao et al. (2020); Lind et al. (2017); McArthur (2016); Narciso et al. (2019); Oh, Herrera, and Bailenson (2019); Pizzoli et al. (2019); Prpa et al. (2017); Rzayev et al. (2019); Sawada et al. (2020); Erik Sikström et al., (2016); Suarez et al. (2017)	22	Presence	Gao, Kim, and Kim (2018); Liao et al. (2020); Wu et al. (2018)	3
Immersion	Davies et al., (2017); Fletcher, Hulusic, and Amelidis (2019); Geronazzo et al., (2019); Gospodarek et al., (2019); Jeon and Jo (2019); Liao et al. (2020); Ooi et al., (2017); Pedersen et al. (2020); Prpa et al. (2017); Rogers et al. (2018); Yan, Wang, and Li (2019)	11	Immersion	Geronazzo et al., (2019); Wu et al. (2018)	2
Perspicuity/usability	Y. Li, Tennent, and Cobb (2019); Michailidis et al., (2019); Erik Sikström et al., (2016)	3	Distraction/disturbance	Bombeke et al. (2018); Souza Silva et al. (2019)	2
Distraction/disturbance	Ghosh et al. (2018); Jeon and Jo (2019)	2	Recall	Byrns et al. (2020); Oh, Herrera, and Bailenson (2019)	2
Task load	Chen, Wu, and Zhu (2018); Moraes et al. (2020)	2	Recognition	Chen, Wu, and Zhu (2018); Oh, Herrera, and Bailenson (2019)	2
Understandability/comprehension	Ghosh et al. (2018); Rzayev et al. (2019)	2	Noticeability	Ghosh et al. (2018)	1
Noticeability	Ghosh et al. (2018)	1	Restorativeness	Li and Xie, (2018)	1
Restorativeness	Park et al. (2020)	1	Understandability/comprehension	Ghosh et al. (2018)	1
Motivational (n = 34)					
Flow	Bombeke et al. (2018); Deacon, Stockman, and Barthet (2017); Dong, Wang, and Guo (2017); Fletcher, Hulusic, and Amelidis (2019); Pedersen et al. (2020); Prpa et al. (2017)	6	Agency/dependability	Prpa et al. (2018)	1
Attentional focus	Bird et al. (2019, 2020); Geronazzo et al., (2018); E. Sikström et al., (2016)	4	Attentional focus	Byrns et al. (2020)	1
Agency/dependability	Y. Li, Tennent, and Cobb (2019); Michailidis et al., (2019); Sikström, de Götzen, and Serafin (2015)	3	Engagement	Bombeke et al. (2018)	1
Engagement	Bialkova and Gisbergen (2017); Prpa et al. (2017); Rogers et al. (2018)	3			
Challenge	Fletcher, Hulusic, and Amelidis (2019); Pedersen et al. (2020)	2			
Competence	Fletcher, Hulusic, and Amelidis (2019); Pedersen et al. (2020)	2			
Efficiency	Y. Li, Tennent, and Cobb (2019); Michailidis et al., (2019)	2			
		2			

(Continued)

Table 8. Continued.

Subjective self-report			Objective (performance/usage/physiological/linguistic)		
Outcome	References	n	Outcome	References	n
Helpfulness of sensory cue	Feng, Dey, and Lindeman (2016); Fletcher, Hulusic, and Amelidis (2019)				
Intent of supportive action	Bialkova and Gisbergen (2017); Shin et al. (2019)	2			
Involvement	Dong, Wang, and Guo (2017); Kruijff et al. (2016)	2			
Performance accuracy	Pedersen et al. (2020)	1			
Recreational value	Lindquist et al., (2020)	1			
Support for creativity	Deacon, Stockman, and Barthet (2017)	1			
Urgency	Ghosh et al. (2018)	1			
Parasocial/interpersonal (n = 10)					
Social presence	Lee, Bruder, and Welch (2017); Oh, Herrera, and Bailenson (2019); Rzayev et al. (2019); Shin et al. (2019)	4	Minimum comfortable distance	R. Li, Tennent, and Cobb (2019)	1
Humanness	Rzayev et al. (2019)	1	Peripersonal space perception		1
Interpersonal liking	Oh, Herrera, and Bailenson (2019)	1	Social presence	Lee, Bruder, and Welch (2017)	1
Parasocial interaction	Shin et al. (2019)	1			
Audio properties (n = 46)					
Loudness	Guang-yu, Yu-an, and Haochen (2019); Hong et al. (2019); Ishikawa, (2019); Jeon and Jo (2019); Jo and Jeon (2020); Maffei et al. (2013); Masullo, Maffei, and Pellegrino (2019); Ong et al., (2017); Schoeffler et al. (2015)	9	Expected sound level	Letwory, Aalmoes, and Miltenburg (2018)	1
Audio quality	Geronazzo et al., (2018); Holm, Väänänen, and Battah (2020); Hong et al., (2018); Ong et al., (2017); Prpa et al. (2017); Wu et al. (2018)	6	Preferred ambience	Park et al. (2020)	1
Distance	Ooi et al., (2017); Schoeffler et al. (2015); Suarez et al. (2017); Yan, Wang, and Li (2019)	4	Preferred ambience level	Lee and Lee (2017)	1
Reverberance	Ishikawa, (2019); Ooi et al., (2017); Schoeffler et al. (2015)	3			
Sound source directionality	Jeon and Jo (2019); Kasuya et al. (2019); Ooi et al., (2017)	3			
Audio model preference	Davies et al., (2017); Hong et al., (2018)	2			
Sound source externalisation	Jeon and Jo (2019); Ooi et al., (2017)	2			
Audio model difference	Hong et al., (2018)	1			
Clarity	Ishikawa, (2019)	1			
Control over sonic interaction	Deacon, Stockman, and Barthet (2017)	1			
Distinctiveness	Ooi et al., (2017)	1			
Dominance of soundscape elements	Ooi et al., (2017)	1			
Duration	Ishikawa, (2019)	1			
Expected sound level	Cranmer et al. (2020)	1			
Imagining of historic soundscape	Falconer et al., (2020)	1			
Loudspeaker accuracy	Yan, Wang, and Li (2019)	1			
Overall listening experience	Schoeffler et al. (2015)	1			
Sound source identification	Shin et al. (2019)	1			
Sound source localisation accuracy	Yan, Wang, and Li (2019)	1			
Sound source width	Jeon and Jo (2019)	1			
Vibrancy	Ooi et al., (2017)	1			
Physiological outcomes (n = 4)					
			Autonomic nervous system (ANS) activity	Bird et al. (2020)	1
			Brain activity	Andreano et al., (2009)	1
			Heart rate	Chen et al. (2017b)	1
			Pupil dilation	Chen et al. (2017b)	1

synchrony, integration of visual elements into the environment, removal of audio/visual stimuli, virtual crowd density, and extent of biodiversity. The latter includes ownership and identification with an avatar, but mostly focuses on specific bodily sensations including awareness of bodily movement, heart rate, and para-sympathetic activity as well as an illusory sense of movement, haptic feedback, and increased weight.

Perhaps unsurprisingly, the most common cognitive measures are that of presence and immersion. Other aspects in this category can be generally related to the ability to understand and interact with virtual environments through mediating technologies, such as general usability/perspicuity, task load, ability to recall information, and understandability or recognition of stimuli for these purposes. Conversely, some measures were focused on distraction or disturbance from tasks or restorativeness gained during respite from performing tasks. The most common outcomes related to motivation were those relating to absorption within tasks, including flow, focus, engagement, and involvement. The ability of audio to aid in task completion was studied through perceived competence, efficiency, sense of task performance accuracy, and helpfulness of stimuli in completing tasks. Measures of a call to action included aspects of the stimuli itself, specifically notice-ability and urgency, enabling factors toward behaviour change, specifically perceived support for creativity and recreational value, and direct behaviour changes in the form of an intent toward financial support.

With the exception of loudness and audio quality, outcomes related to audio properties most often related to the spatial perception of audio, such as sound source distance, width, reverberance, perception as coming from a specific direction or as from originating external to one's head, and a sense of accuracy in identifying the spatial position of audio sources. Another common audio-related focus was on listener preference, such as preference for specific ambiences, ambience levels, and different models for presenting audio. The least-studied experiential outcome category was that of parasocial/interpersonal experiences which could generally be grouped into outcomes related to the general perception of other social beings, i.e. social presence, parasocial interaction, and peripersonal space perception, and general perception or affective social measures, i.e. perceived humanness, interpersonal liking, and the minimum comfortable distance with robots. Finally, some physiological measures were not explicitly linked to experiential outcomes but were reported as is since it could be argued that they are still directly related to specific experiences.

As mentioned in section 3.5 above, there were 79 instances of existing questionnaires being used, which includes some questionnaires being used across several different studies. In total there were 49 different questionnaires used. The full list of questionnaires used and how often each was used is given in Table 9. Again, items with the same number of instances used are listed together for brevity.

Perhaps the most notable disparity between the frequency of outcomes measured and that of questionnaires used is that while simulator sickness ($n = 7$) was not as commonly measured as several other experiential outcomes, the simulator sickness questionnaire (SSQ) was one of the most commonly used questionnaires. This can be attributed to the relatively standardised way to measure simulator sickness, especially when compared to measuring other outcomes such as presence or positive affect which were measured with eight and seven different existing questionnaires respectively, in addition to those created for individual studies. In contrast, none of the questionnaires listed was used to study perceived pleasantness; in other words, although pleasantness was one of the most-commonly investigated outcomes, it was exclusively studied with questionnaires created for the purpose of each study. The same applies to outcomes for Audio Properties, with the exception of the use of the magnitude estimation method for perceived loudness (Ong et al. 2017), spatial audio quality inventory for perceived quality (SAQI) (Geronazzo et al. 2018), overall listening experience

Table 9. Questionnaires used.

Questionnaire	n
Simulator sickness questionnaire (SSQ), Igroup presence questionnaire (IPQ)	5
Witmer and Singer questionnaire, Positive and negative affect scale (PANAS), Subjective unit of discomfort scale (SUD), game experience questionnaire (GEQ), Slater Usoh Steed questionnaire (SUS), self-assessment manikin (SAM)	3
NASA TLX questionnaire, State-trait anxiety inventory (STAI), ITC sense of presence inventory (ITC-SOPI), user experience questionnaire (UEQ), feeling scale (FS), felt arousal scale (FAS), Borg CR10 scale, physical activity enjoyment scale (PACES), Tammen's attentional scale, DBT® diary cards	2
Fast motion sickness scale (FMS), Attitude Towards Public Speaking Questionnaire (ATPSQ), uncanny valley effect questionnaire, discrete emotions questionnaire (DEQ), subjective units of distress scale (SUD), numerical pain rating scale, magnitude estimation (ME) method, spatial audio quality inventory (SAQI), overall listening experience (OLE) questionnaire, spatial presence experience scale (SPES), presence involvement flow framework 2 (PIFF2), Swedish Viewer-User Presence Questionnaire (SVUP), spatial perception questionnaire (SPQ), immersive response questionnaire (IMX), immersed tendency questionnaire (ITQ), immersive experience questionnaire (IEQ), presence questionnaire (PQ), Perceived Restorativeness Soundscape Scale (PRSS), biodiversity experience index (BEI), ownership + agency + threat questionnaire, player identification scale (PIS), social presence questionnaire, co-presence questionnaire, Temple Presence Inventory, Parasocial Interaction Scale, dispositional positive emotions scale (DPET), brief awe scale, game engagement questionnaire (GEngQ), creativity support index (CSI), flow questionnaire (FQ), enjoyment scale[a]	1

[a]Based on adjective items for enjoyment extracted from (Tamborini et al. 2010).

Table 10. The effects of audio on user/player experience organised by categories related to experience.

	Effects		
Measured outcome	Negative effect	No effect	Positive effect
Emotion/affect			
Pleasantness	Echevarria Sanchez et al. (2017a); Hong et al. (2019); Jo and Jeon (2020); Masullo, Maffei, and Pellegrino (2019)		Echevarria Sanchez et al., (2017a); Hong et al. (2019); Jo and Jeon (2020); Kelling et al. (2018); Park et al. (2020)
Discomfort	Jo and Jeon (2020); North et al. (2008)	Peng et al. (2020)	Aalmoes and Zon (2019); Brinkman, Hoekstra, and van Egmond (2015); Jo and Jeon (2020); Josman et al. (2008)
Positive affect	Bird et al. (2020); Park et al. (2020)	Chen et al. (2017b); Chirico and Gaggioli (2019)	Bird et al. (2019); Park et al. (2020); Pizzoli et al. (2019)
Activation/arousal	Pizzoli et al. (2019)	Chen et al. (2017b)	Bird et al. (2019); Chen et al. (2017a)
Enjoyment		Bird et al. (2019, 2020); Erik Sikström et al., (2016)	
Simulator sickness		Feng, Dey, and Lindeman (2016); Peng et al. (2020); Sawada et al. (2020)	
Fear/anxiety	Johnston, Egermann, and Kearney (2020)	Lee, Bruder, and Welch (2017)	Brinkman, Hoekstra, and van Egmond (2015)
Tension/annoyance			Jeon and Jo (2019); Maffei et al. (2013); White et al. (2014)
Calmness	Jo and Jeon (2020)		Jo and Jeon (2020)
Dominance		Chen et al. (2017b); Pizzoli et al. (2019)	
Exertion		Bird et al. (2019, 2020)	
Liking/affective attraction	Bialkova and Gisbergen (2017)	Lee, Bruder, and Welch (2017)	
Negative affect		Chen et al. (2017b); Chirico and Gaggioli (2019)	
Affective valence		Oh, Herrera, and Bailenson (2019)[b]	
Awe		Chirico and Gaggioli (2019)	
Eventfulness			Jo and Jeon (2020)
Frustration	Byrns et al. (2020)		
Interest			Kelling et al. (2018)
Pain intensity	Zeroth, Dahlquist, and Foxen-Craft (2019)		
Pain tolerance		Zeroth, Dahlquist, and Foxen-Craft (2019)[b]	
Relaxation			Pizzoli et al. (2019)
Satisfaction		Oh, Herrera, and Bailenson (2019)[b]	
Stimulation			Michailidis et al., (2019)
Perceptual/embodiment			
Realism			Feng, Dey, and Lindeman (2016); Huang et al., (2019); Jeon and Jo (2019); Lindquist et al., (2020); Peng et al. (2020); Sawada et al. (2020)
Preference		Park et al. (2020)	Chen et al. (2017b); Lindquist et al., (2020); Park et al. (2020)
Heart rate awareness			Chen et al. (2017b)
Sense/awareness of body movement			Kruijff et al. (2016); Feng, Dey, and Lindeman (2016)[a]
Believability			Gao, Kim, and Kim (2018)
Biodiversity			Lindquist et al., (2020)
Crowd density			Jo and Jeon (2020)
Naturalness		Bialkova and Gisbergen (2017)	
Ownership			Sikström, de Götzen, and Serafin (2015)[a]
Parasympathetic activity		Bird et al. (2020)	
Cognitive			
Presence		Bialkova and Gisbergen (2017); Feng, Dey, and Lindeman (2016); Sawada et al. (2020); Wu et al. (2018); Oh, Herrera, and Bailenson (2019)[b]	Brinkman, Hoekstra, and van Egmond (2015); Lee, Bruder, and Welch (2017); Liao et al. (2020)[a]
Attentional focus		Bird et al. (2020)	Bird et al. (2019); Byrns et al. (2020)[a]
Understandability/ comprehension			Ghosh et al. (2018)[a]
Recognition		Oh, Herrera, and Bailenson (2019)[b]	
Perspicuity/usability		Erik Sikström et al., (2016)	
Recall	Oh, Herrera, and Bailenson (2019)[b]		Byrns et al. (2020)
Distraction/ disturbance			Ghosh et al. (2018); Jeon and Jo (2019)
Immersion			Jeon and Jo (2019)
Motivational			
Recreational value		Lindquist et al., (2020)	Lindquist et al., (2020)

(*Continued*)

Table 10. Continued.

Measured outcome	Effects		
	Negative effect	No effect	Positive effect
Engagement		Bialkova and Gisbergen (2017)	
Helpfulness of sensory cue			Feng, Dey, and Lindeman (2016)
Intent of supportive action			Bialkova and Gisbergen (2017)
Involvement			Kruijff et al. (2016)
Noticeability			Ghosh et al. (2018)
Performance accuracy			Pedersen et al. (2020)
Urgency			Ghosh et al. (2018)
Parasocial/interpersonal			
Interpersonal liking		Oh, Herrera, and Bailenson (2019)[b]	
Minimum comfortable distance	R. Li, Tennent, and Cobb (2019)[a]		
Social presence	Oh, Herrera, and Bailenson (2019)[b]		
Audio properties			
Loudness			Jeon and Jo (2019); Masullo, Maffei, and Pellegrino (2019)
Expected real-world sound level	Cranmer et al. (2020)		
Imagining of historic soundscape			Falconer et al., (2020)
Physiological outcomes			
Autonomic nervous system activity		Bird et al. (2020)	
Brain activity			Baka et al., (2018)
Heart rate		Chen et al. (2017b)	
Pupil dilation		Chen et al. (2017b)	

[a]Partial support for hypothesis.
[b]Extraneous audio sources, i.e. outside of VR.

questionnaire for perceived overall listening experience (Schoeffler et al. 2015), and Witmer and Singer questionnaire for perceived sound source identification (Shin et al. 2019).

It should also be noted that most questionnaires that are used twice are used in two related studies by one/more of the same author. Specifically, this includes the feeling scale (FS), felt arousal scale (FAS), Borg CR10 scale, physical activity enjoyment scale (PACES), and Tammen's attentional scale used in (Bird et al. 2019, 2020) and the use of DBT® diary cards in (Gomez et al. 2017; Nararro-Haro et al. 2016). As such, of the 49 different questionnaires used, only the following 12 were used more than once in unrelated studies: Simulator sickness questionnaire (SSQ), Igroup presence questionnaire (IPQ), Witmer and Singer questionnaire, Positive and negative affect scale (PANAS), Subjective unit of discomfort scale (SUD), game experience questionnaire (GEQ), Slater Usoh Steed questionnaire (SUS), self-assessment manikin (SAM), NASA TLX questionnaire, State-trait anxiety inventory (STAI), ITC sense of presence inventory (ITC-SOPI), and user experience questionnaire (UEQ).

3.8. Study outcomes – effects of audio

The list of measured outcomes and how each was affected by the introduction of audio is listed in Table 10. Similar to Table 8 above, the list only includes discretely measurable outcomes and thus does not include qualitative data. Positive and negative effects indicate that a measured outcome was significantly increased or decreased with the introduction of audio stimuli. Furthermore, it should be noted that these results only represent instances where it is possible to note the effect of the introduction of audio on its own as opposed to instances where the differences between factors such as types/implementations of audio, properties of participants, etc. are being studied and that case studies are excluded.

One notable finding is that some experiential outcomes are simultaneously affected in seemingly contradictory ways, specifically in terms of experiential outcomes that are positively and negatively affected by audio in the same study, such as pleasantness (Echevarria Sanchez et al. 2017a; Hong et al. 2019; Jo and Jeon 2020), positive affect (Park et al. 2020), and discomfort, relaxation, and calmness (Jo and Jeon 2020). This can be attributed to the fact that all these studies fall within the domain of soundscape studies and all study the differences between natural and artificial/human ambiences. For each case listed here, natural ambience is found to increase positive affect (pleasantness, relaxation, calmness) and decrease negative affect (discomfort) and artificial/human ambience has the inverse effect. Other soundscape research also reports natural ambiences as

increasing pleasantness or restorativeness and artificial/ human ambiences as decreasing pleasantness and increasing anxiety effects (Li and Xie 2018; Ong et al. 2017; Ooi et al. 2017). Related to this (Lindquist et al. 2020), studied the experiential differences between generic ambiences, hich had no effect on perceived recreational value, and detailed ambience created for specific visuals, leading to increased perceived recreational value.

The large variability in study outcomes for emotional/affective measures can be partially explained by the differing use of audio in this area, for example, using it for short-term increases in discomfort or fear/anxiety (Brinkman, Hoekstra, and van Egmond 2015; Josman et al. 2008) as well as for decreasing them with repeated exposure (Johnston, Egermann, and Kearney 2020; North et al. 2008). While one study (Lee, Bruder, and Welch 2017) found no effect of audio on fear/anxiety, it should be noted that the base rates for negative affect was perceived as relatively high, which might explain the lack of a significant change. Another example of seemingly contradicting findings is using audio to reduce arousal/stimulation and increase calmness (Pizzoli et al. 2019), such as for meditation practices, or have the opposite effect, such as for entertainment or exercise purposes (Bird et al. 2019; Chen et al. 2017a). There are also some notable outcomes where the introduction of audio did not have an effect, such as enjoyment, awe, and the relatively large number of studies on presence. Finally, there is also a shortage of research investigating the role of audio in parasocial/interpersonal experiences.

4. Discussion

This paper has reviewed the literature regarding how audio is used to affect aspects of the user/player experience in VR in various domains.

4.1. Domain areas of concern

This section serves as an answer to RQ1 (In which contexts are the effects of audio in HMD-based VR being studied?). For the studies that did focus on a specific domain area, the majority focused on entertainment, soundscape research, and therapy/rehabilitation while the largest focus was on medium-specific studies of VR itself. Within this discussion, however, the potential effects of a bias introduced by the choice of databases consulted must be acknowledged as a potentially biasing factor.

4.1.1. Entertainment focuses on new approaches for music but does not represent the size of the games market

The corpus shows that the effects of audio are commonly studied in the domains of entertainment (19 studies), soundscape research (18 studies), and therapy (18 studies). The research in entertainment generally focused on music, with 10 of the 19 studies in this category enabling ways of music listening, training, or performing (Chang, Kim, and Kim 2018; Deacon, Stockman, and Barthet 2017; Droste, Letellier, and Sieck 2018; Fletcher, Hulusic, and Amelidis 2019; Holm, Väänänen, and Battah 2020; Kasuya et al. 2019; Lind et al. 2017; Pedersen et al. 2020; Rees-Jones and Daffern 2019; Shin et al. 2019). The corpus also contains only three studies focusing on games (Bombeke et al. 2018; Rogers et al. 2018; Wedoff et al. 2019) which is surprising given not only the current size of the market for VR games (Grand View Research 2020) but also considering that most of the studies were published between 2017 and 2020, by which point the VR games market was already worth several million USD (Statista 2022). The role of audio in games is often noted as being underutilised (Ekman 2005; Ribeiro et al. 2020) and research is disproportionately slanted toward certain genres, especially horror games (Ribeiro et al. 2020; Rogers et al. 2018). While the influence of our choice of databases must be acknowledged as a potential influence on the representation of this field in the corpus, there is still a notable discrepancy between the size of the market (as a proxy measure for interest in the field) and the number of studies focusing on audio.

4.1.2. Soundscapes research emphasizes the dichotomy between artificial and natural soundscapes but lacks a validated measure for pleasantness

Soundscape research generally made use of VR to reproduce urban/architectural environments and study the effects of noise (Aalmoes and Zon 2019; Cranmer et al. 2020; Echevarria Sanchez et al., 2017a, 2017b; Guang-yu, Yu-an, and Haochen 2019; Jeon and Jo 2019; Letwory, Aalmoes, and Miltenburg 2018; Maffei et al. 2013; Masullo, Maffei, and Pellegrino 2019; White et al. 2014) and generally confirm the utility of using VR for this purpose. The large representation of soundscape studies in the corpus is not surprising given the focus on audio in the search terms. This is also the area where the most consistent and generalisable findings can be found, i.e. natural ambiences are associated with positive affect measures, most commonly pleasantness (Echevarria Sanchez et al. 2017a; Hong et al. 2019; Jo and Jeon 2020; Park et al. 2020),

while artificial/human ambiences are associated with negative affect measures, such as discomfort and tension/annoyance (Jeon and Jo 2019; Maffei et al. 2013; White et al. 2014). While the domain of soundscape research has seen an argument has been made against an overly simplistic dichotomous categorisation of natural and artificial ambiences (Harriet and Murphy 2015) the corpus not only suggests that this dichotomy is still common in VR soundscape research, but also confirms the pleasant/restorative effects of natural ambiences (Echevarria Sanchez et al. 2017a; Hong et al. 2019; Jo and Jeon 2020; Park et al. 2020) and opposite effect for artificial/human ambiences (Jeon and Jo 2019; Maffei et al. 2013; White et al. 2014).

Previous research has suggested that increased immersion would be beneficial to this field and suggests VR as one possible tool to improve soundscape research (Erfanian et al. 2019). While the comparison of different approaches, e.g. VR vs. in-situ, was out of the scope of the current study, the fact that the results for soundscape studies in VR agree with the findings of in-situ studies, e.g. the restorativeness of natural ambience and the opposite effect of artificial ambience (Uebel et al. 2021) provides evidence for the usefulness of VR as a tool for soundscape appraisal and investigation into the psychological effects of different sound environments. However, this field of research also highlights what is perhaps the largest methodological shortcoming in the corpus, which is the lack of a pleasantness questionnaire, as discussed in section 4.3.1 below.

4.1.3. Audio is effective in exposure therapy, but the focus is primarily on passive listening

The two therapeutic areas where VR was most common were exposure therapy (7 studies) and mindfulness practice (6 studies). The use of audio was found to be effective both in terms of creating the experience of anxiety in the short term (Brinkman, Hoekstra, and van Egmond 2015; Josman et al. 2008) and in reducing it through repeated exposure (Johnston, Egermann, and Kearney 2020; North et al. 2008), which provides supporting evidence for the effectiveness of audio in a virtual reality exposure therapy (VRET) context. There was also some support for the use of audio in a mindfulness context to increase relaxation and decrease activation/arousal (Gomez et al. 2017; Nararro-Haro et al. 2016; Pizzoli et al. 2019; Prpa et al. 2017, 2018; Seabrook et al. 2020), although the bulk of these results focused on qualitative data which means that the reporting and analysis of these results was out of the scope of the current study; these might be explored further in other types of review papers, such as narrative reviews. Other studies highlight the potential for audio to aid

in various therapeutic outcomes, such as using music therapy to reduce frustration and increase recall for Alzheimer patients (Byrns et al. 2020), as well as using audio as a distraction to reduce pain intensity (Zeroth, Dahlquist, and Foxen-Craft 2019). While these findings highlight the importance of audio in creating the desired experiential outcomes, further investigation into these studies shows that they are overwhelmingly focused on passive listening rather than active participation in the creation of audio. The implications of this are not clear, since both active participation and passive listening have been linked to improvements in mental well-being (Rogers and Nacke 2017), although active participation affords other potential benefits such as social interactions (a category of experiences which is severely lacking in the corpus). Examples of other audio-focused phenomena that present the potential to be useful in a therapeutic context include binaural beats (Rahman et al. 2021) or autonomous sensory meridian response (ASMR) (Fredborg et al. 2021); the effects of these in VR, however, have not been studied.

4.1.4. Accessibility highlights the promise of VR to create audio-only experiences for those with visual impairments

While relatively sparse, the six accessibility-focused studies show some evidence of the potential utility for VR as an accessibility tool (Bălan et al. 2018; Dong, Wang, and Guo 2017; Picinali et al. 2014; Siu et al. 2020; Souza Silva et al. 2019; Wedoff et al. 2019). Four of these studies specifically make use of 'audio-only' VR that makes use of head-tracking to control audio feedback (Bălan et al. 2018; Picinali et al. 2014; Siu et al. 2020; Wedoff et al. 2019) which bears some resemblance to audio-based games that deliver information about the virtual environment using audio, either primarily, e.g. Beowulf (Liljedahl, Papworth, and Lindberg 2007) or exclusively, e.g. Papa Sangre (Collins and Kapralos 2012). In traditional gaming contexts, these games demonstrate the effectiveness of audio as a tool to convey spatial information and to create immersive gameplay experiences for those with visual impairments (Collins and Kapralos 2012). These games, however, are still quite limited compared to the affordances of VR technology. For example, Papa Sangre, which is a mobile game, utilises the built-in accelerometer of the mobile device, allowing the player to 'look around' by moving their device. While novel, the underlying technology for this implementation is still comparatively limited as it cannot provide the accuracy nor intuitiveness of moving one's own head relative to the environment. The head-tracking technology of VR devices thus provides opportunities to utilise the designs of these

existing games in a variety of contexts, including games. In a more theoretical sense, the use of such audio-only solutions also provides the opportunity to study the effect of audio in HMD-based VR completely separate from the influence of visual feedback. However, the six studies listed either focus on qualitative data or comparisons of audio types or participant characteristics, which was not in the scope of this study. Further research in accessible VR, specifically for those with visual impairments, might thus consider the effect of such systems with more focused systematic reviews.

4.1.5. Arts and cultural heritage primarily focus on recreating museum experiences although other possibilities are being explored

Out of nine studies exploring the use of VR in some context related to arts and cultural heritage, 6 of these focused on creating a virtual version of a traditional, i.e. brick-and-mortar museum (Bialkova and Gisbergen 2017; Kelling et al. 2018; Li, Tennent, and Cobb 2019; Rzayev et al. 2019; Sprung et al. 2018; Tennent et al. 2020). The implications for the use of sound in these contexts are unclear with music having both a positive effect on pleasantness (Kelling et al. 2018) and a negative effect on liking/affective attraction (Bialkova and Gisbergen 2017). Most of these studies also focused on the use of qualitative data, which was out of the scope of the current study. Rather than recreating museum spaces, one study focused on the use of VR to create an archeological space (Falconer 2017) and two studies explored a completely different use of VR technology in creating an art installation (Prpa et al. 2017, 2018). While the use of VR in contexts for creating art installations and exhibitions is receiving increased attention (Kim and Lee 2022; Parker and Saker 2020), the effect of VR within the experience of these contexts is still comparatively unexplored. Furthermore, the focus remains on recreating existing spaces rather than exploring new possibilities afforded by VR technologies, including the use of audio feedback.

4.1.6. VR as a medium is rapidly growing but mass adoption is still far away

The largest group of studies, however, did not fall into a clear domain area but rather investigated the use of audio in VR or, more broadly, the medium of VR itself (Çamci 2019; Chen et al. 2017a, 2017b, 2018; Chirico and Gaggioli 2019; Chittaro 2012; Feng, Dey, and Lindeman 2016; Gao, Kim, and Kim 2018; Ghosh et al. 2018; Kruijff et al. 2016; Lee, Bruder, and Welch 2017; Lee and Lee 2017; Liao et al. 2020; Narciso et al. 2019; Oh, Herrera, and Bailenson 2019; O'Hagan, Williamson, and Khamis 2020; Peng et al. 2020; Sawada et al. 2020;

Shimamura et al. 2020; Sikström et al. 2015, 2016a, 2016b; Van den Broeck, Pamplona, and Fernandez Langa 2017; Zhang et al. 2018; Zhao et al. 2017). This serves as an indication that, while the medium of VR has seen substantial growth, there is still much uncertainty for both researchers and practitioners as to how this medium should be used and studied. There is still a common view that VR technology is in its infancy (Radianti et al. 2020; Zaini et al. 2022) and the large number of studies that do not focus on VR in a specific domain area but rather investigate audio or VR at large suggests that mass adoption of VR technologies in domain areas is yet to be achieved.

4.2. Notable effects of audio

This section provides answers to RQ2 (Which aspects of experience are being studied with regards to audio in HMD-based VR research?) and RQ3 (How does audio affect various aspects of experience in HMD-based VR?). The aspects of experience studied most frequently are presence, realism, immersion, and pleasantness. The effect of audio on pleasantness is one of the most consistent findings across the corpus, i.e. natural or artificial soundscapes are respectively perceived as pleasant or unpleasant. The effect of audio on realism is also very consistent, although perhaps unsurprising: the introduction of audio increases perceived realism. What is notable here, however, is the narrow focus on verisimilitude rather than perspectives on realism that provide for unreal or impossible scenarios. Perhaps more surprisingly, the sense of presence is mostly unaffected by the introduction of audio and only one study investigated the effect of the introduction of audio on its own on the perception of immersion. Furthermore, it must be acknowledged that the search string used imposes a necessary requirement on certain keywords, which could have affected the results obtained.

4.2.1. Presence remains the most important measure in VR studies but is often unaffected by audio along with enjoyment and awe

The fact that presence is the most-studied outcome in the corpus ($n = 25$) concurs with the long- and widely held stance that presence is one of the most important aspects of a VR experience (Heineken and Schulte 2007; Ijsselsteijn et al. 2000; Kern and Ellermeier 2020; Steuer 1992). The term, however, is not without contention (Slater 2009) and, as pointed out by (Slater 2003) the meaning in VR research is often confused or equated with that of immersion. There is also a large number of results showing a lack of increase in certain outcomes, including presence (Bialkova and Gisbergen 2017; Feng,

Dey, and Lindeman 2016; Oh, Herrera, and Bailenson 2019; Sawada et al. 2020; Wu et al. 2018) but also enjoyment (Bird et al. 2019, 2020; Sikström et al. 2016a) and awe (Chirico and Gaggioli 2019) which is perhaps surprising given the emphasis on audio in entertainment (Collins 2020; Kauhanen et al. 2017) and the general view emphasizing the role of audio in perceiving space (Larsson et al. 2010; Nordahl and Nilsson 2014). It is important to note that each of the studies on awe and enjoyment made use of music and used different processes for selecting music, which might have different effects based on the preferences of participants (Chirico and Gaggioli 2019). More rigorous protocols for audio selection, especially music, would thus be a valuable contribution to improve the validity of audio studies in VR.

4.2.2. VR experiences are frequently evaluated in terms of verisimilitude with real environments

One of the most frequently measured outcomes was realism ($n = 15$), i.e. the level of verisimilitude between some aspect of the VR experience and its counterpart in physical reality. Audio is found to consistently increase the sense of realism, which highlights the importance of audio in contexts where realism is a desirable outcome. However, this perspective of realism as verisimilitude also presents several flaws, such as the fact that what we perceive to be realistic often deviates from reality, e.g. sound design for film and games, i.e. foley often creates sound effects from completely unrelated source materials, such as using audio tape that is bunched together to create footstep sounds (Garner 2017, 76). Users of VR are also able to attribute a sense of realism and adapt their sensory responses accordingly to physically impossible scenarios, such as experiencing a sense of touch from blocks that float in mid-air (Bosman, de Beer, and Bothma 2021). Perhaps more importantly, however, is the fact that VR is not limited to recreating reality, but can create entirely fictional experiences, which limits the usefulness of realism as a construct (Bergstrom et al. 2017; Rogers et al. 2022). As such, several alternatives to realism have been proposed, including credibility (Bouvier 2008), authenticity (Gilbert 2016), plausibility (Bergstrom et al. 2017), and believability (Togelius et al. 2012); for an overview of how the term 'realism' is used across various media/domains, the reader is referred to (Rogers et al. 2022). However, the corpus shows that assessing realism is still a common outcome within the category of outcomes related to perception and/or embodiment within VEs, although there are a few instances of using alternatives such as naturalness (Bialkova and Gisbergen 2017; Chang, Kim, and Kim 2018; Shimamura et al. 2020;

Sikström et al. 2016b) or believability (Gao, Kim, and Kim 2018).

4.2.3. Parasocial/interpersonal experiences are underexplored

With regards to the types of experience being studied, there is a shortage of research investigating the role of audio in parasocial/interpersonal experiences. Given the recent significant increase in using remote collaboration not only for work but for all manner of social interactions, due largely to the COVID-19 pandemic (Choi and Diehm 2021), VR has gained new relevance as a medium for meaningful social interaction (Choi and Diehm 2021; Lee, Joo, and Jun 2021), including the recent resurgence of interest in the concept of immersive, interconnected, virtual social spaces known as the Metaverse (Mystakidis 2022). Previous work in social VR has investigated the role of haptic feedback (Hoppe et al. 2020), gaze-input (Seele et al. 2017), and biofeedback (Desnoyers-Stewart et al. 2019). A variety of audio properties, including verbal and non-verbal, play a key role in the general experience of social interactions and social environments (Poudyal et al. 2021) and social interaction, in turn, plays a key role in some audio-driven practices such as music (D'Ausilio et al. 2015). Given the importance of audio properties, such as being able to localise audio for social interaction (Wie, Pripp, and Tvete 2010), it thus follows that an investigation into the range of audio possibilities and their effects on social experiences would be beneficial towards guiding the design of audio in social VR applications.

4.3. Methodological concerns

This section provides a preliminary answer to RQ4 (What implications from the literature can be drawn regarding the design of future HMD-based VR audio research?). There is a general reliance on subjective self-reports in the corpus, although the fact that many of these were developed for the purposes of the study might bring the validity of the results from these instruments into question. Novelty bias is also noted as a likely confounding variable, although the possibility of simulator sickness makes the implications for this less clear. The vast majority of studies were laboratory-based and made use of experimental or crossover designs, which leads to a lack of descriptive research that draws on natural settings. Finally, there were some instances of using alternative data collection and analysis methods, which might prove beneficial for measuring especially the most common outcomes, such as presence. The

implications for future research are discussed in depth in section 5.

4.3.1. Most studies relied on subjective self-reports but often used questionnaires created for the study

While pleasantness is one of the most common outcomes in the realm of soundscape studies, none of the studies investigating perceived pleasantness used a validated questionnaire for this purpose, which might bring the validity and generalizability of these results into question. The creation and validation of a pleasantness questionnaire would thus be an invaluable contribution toward the domain of soundscape research. Similarly, most measured outcomes in the 'Audio Properties' category listed were measured using a questionnaire created for the purposes of each study. Furthermore, the frequent occurrence of questionnaires created for the purposes of individual studies might impact the validity of results gleaned from these instruments, since they are generally not validated and often rely on single-item scales. However, given the large variety of outcomes from the corpus, a broad standardisation of instruments across studies would require a more focused analysis that considers types and roles of audio in this context.

There are some outcomes with more standardised instruments of measurement, such as the use of the simulator sickness questionnaire (SSQ). While this standardisation can be considered beneficial for the generalizability of such results obtained, there has been criticism against the use of the SSQ in HMD-based VR (Hirzle et al. 2021). It should also be noted that similar criticisms have been raised against the game experience questionnaire (GEQ) (Law, Brühlmann, and Mekler 2018) and against how the measuring of presence is approached in VR research (Slater 2009) with the suggestion that subjective and objective measures should be combined (Ijsselsteijn et al. 2000). This argument is also supported by evidence of the disconnect between self-reported predictions of behaviour and actual behaviours in VR (Slater et al. 2020). While such a combined approach is used sporadically to measure presence and social presence, the overwhelming majority of studies rely primarily on subjective self-reports for presence measurement.

4.3.2. Novelty bias is a likely confounding variable

Another possible reason for the lack of increase in certain outcomes with the introduction of audio is that of novelty bias of using a HMD to experience VR, since the experience might initially be new or unusual to many participants after which they acclimatise with repeated exposure, thus leading to unexpected results. Several of the included studies note this as a potential confounding factor, e.g. (Lind et al. 2017; McArthur 2016; Moraes et al. 2020; Oberman, Bojanić Obad Šćitaroci, and Jambrošić 2018; Rogers et al. 2018; Sikström et al. 2016a; Suarez et al. 2017; Vosmeer and Schouten 2017). The effect of this on the measured outcomes connected to audio feedback is expected to be more pronounced than for visual feedback since the experience of mediated binaural/spatial audio, which is the standard for many VR experiences, is expected to be novel to many participants within the context of media experiences (Lind et al. 2017). This might limit participants' conceptual framework to understand and express their experiences connected to such sensory feedback within these types of experiences (Garner 2017, 38).

The impact of initial novelty/unfamiliarity which dissipates throughout the course of data collection is especially noteworthy when considering the large number of crossover studies where the same participants were exposed to two or more experimental conditions. Furthermore, given the highly subjective responses to different stimuli such as music (Rogers et al. 2018; Schoeffler et al. 2015) and the ability of differing technologies such as headphones to influence factors related to experience (Larsson et al. 2010), this also likely explains some of the variability. This emphasizes the need for thorough familiarisation tasks prior to data collection to reduce the effects of this bias (Moraes et al. 2020; Suarez et al. 2017) as well as to consider participants' level of experience with VR (Sikström et al. 2016b; Vosmeer and Schouten 2017). However, the implications this carries for future research is also made more complex by the possibility of inducing simulator sickness that might also introduce a confounding factor with prolonged exposure (Bird et al. 2019).

4.3.3. There is an overall lack of descriptive research

The majority of study methodologies involved exposing participants to some type of technological intervention. This means that there is a general lack of descriptive research that draws from the large VR user/player base, which points to a lack of insight regarding how these technologies are used and experienced outside of laboratory/exhibition settings. Previous descriptive research in, for example, game music has revealed insights that might challenge designers' assumptions regarding player habits regarding in-game audio, e.g. (Rogers and Weber 2019). The overrepresentation of research in lab environments versus natural settings also has negative implications for ecological validity (Kong et al. 2020). Given the large market for consumer VR hardware (Kugler 2021; Baker 2022), similar research would be feasible and might provide useful

insights and design guidelines for VR audio. As an example, the impact of different audio types and roles (sound effects/music, diegetic/non-diegetic, etc.) can be investigated through descriptive research to provide practitioner design guidelines. While guidelines and practitioner examples do exist for some audio properties, such as for audio diegesis (Phillips 2018; Stein and Goodwin 2019), these tend to be speculative and would benefit from empirical research contextualised in their respective domain areas.

4.3.4. Alternative data collection and analysis methods present a promising supplement to subjective self-reports

The corpus also includes a few instances of exploratory research which makes use of alternative data collection methods such as collaborative workshops (Ghosh et al. 2018) as well as data analysis methods such as linguistic inquiry word count (Oh, Herrera, and Bailenson 2019), verbal protocol analysis (Olko et al. 2017), and interaction analysis (Deacon, Stockman, and Barthet 2017). Given the dynamic nature of interactive VR experiences, there has been criticism against using conventional techniques to measure audio quality in these contexts (Yan, Wang, and Li 2019) and alternative approaches have been employed that incorporate varying degrees of interactivity (Robotham et al. 2022). Alternative analysis techniques have been explored in game studies to measure experiential outcomes, such as using grounded theory to measure immersion (Farkas et al. 2020) and facial expression analysis to measure emotions (Roohi et al. 2018). In VR research there has been previous work investigating the use of linguistic analysis to measure presence (Kramer, Oh, and Fussell 2006) as well as audio analysis to measure valence and arousal (Alzayat, Hancock, and Nacenta 2019). This suggests that such approaches could also provide fruitful alternatives for collecting and analyzing experience data in audio-driven VR research. Future research could thus employ more such approaches, possibly alongside more standardised approaches such as subjective-self reports.

5. Future agenda

The previous section presented a discussion of the results from the corpus within the wider contexts of various domain areas of study and pointed out gaps in these areas. This section summarises these gaps in the form of future agendas for research on the effects of audio in HMD-based VR. By providing a thematic and methodological agenda for future audio-focused research in VR, we provide an answer to RQ4 (What implications from the literature can be drawn regarding the design of future HMD-based VR audio research?).

5.1. Thematic agenda

While it could be argued that audio can benefit any of the domain areas listed in Table 3, including those with very few studies, the most notable discrepancy in terms of a domain that is receiving much attention commercially but not in research is that of games. While there is no shortage of previous work emphasizing the importance of audio in digital games at large (Collins 2020), digital games research seems slow to catch up within the audio modality in particular. There is also a severe lack of investigation into the use of audio in para-social/interpersonal VR, which is proving to be a growing area of interest due to the resurgence of the Metaverse concept which promotes multi-user VR across interconnected virtual environments (Mystakidis 2022). It should also be restated that this is not limited to verbal sounds but may include any number of other non-verbal sounds such as music and environmental sounds (Poudyal et al. 2021). Further research is thus needed to understand the role of audio in both of these rapidly growing markets.

While there is more research in the broader domain of therapeutic studies, future research in this area can address the lack of studies that make use of interactive audio. The capabilities of interactive media afford user participation with the audio stimuli being received, which creates opportunities for co-creation rather than passive consumption of such stimuli (van Elferen 2020). One notable benefit of such co-creative affordances in the therapeutic domain is the potential to create social experiences, which might be especially suitable to combat the social isolation brought on by the COVID-19 pandemic (Choi and Diehm 2021). The audio domain at large has also seen an uptake in investigating other audio-focused phenomena, such as binaural beats and ASMR, which could prove useful in therapeutic or mindfulness contexts; future research might investigate the use of such practices in VR-driven therapy. In terms of arts and cultural heritage, the small number of studies focused primarily on recreating existing spaces for art exhibits, which does not realise the full potential of immersive VR in reimagining how arts and heritage may be experienced. Future research might utilise interdisciplinary approaches to designing art exhibitions and experiences and a logical starting point would be a systematic review on the design of such installations and exhibitions.

Lastly, the corpus shows some interest in the use of audio-only VR solutions that do not make use of any

visual feedback and use head-tracking to control audio stimuli. Such implementations provide rich potential for the creation of experiences for those with visual impairments, similar to existing audio-based games like Papa Sangre. Outside of research into accessible technology use, this type of implementation also offers a unique glimpse into the effect of audio feedback in VR without the influence of visual feedback. However, the corpus in this domain was focused on qualitative data or comparison between different characteristics of audio or participants. Future review papers might thus investigate the effect of these implementations in a more focused manner.

5.2. Methodological agenda

The most pressing methodological shortcoming identified in the corpus was the lack of a questionnaire to measure pleasantness, which was the fourth most frequently measured experiential outcome. Across 12 studies that studied perceived pleasantness, each used a questionnaire created for the purposes of the study. The creation of a validated questionnaire for this purpose would thus be an invaluable addition to the domain of soundscape research and possibly the study of audio effects in HMD-based VR at large. A starting point for such an endeavour might be to systematically analyze the existing questionnaires for measuring pleasantness across the corpus as well as to refer to validated instruments for related outcomes found in the corpus such as positive affect (Bradley and Lang 1994; Harmon-Jones, Bastian, and Harmon-Jones 2016; IJsselsteijn, de Kort, and Poels 2013; Watson, Clark, and Tellegen 1988), activation/arousal (Bradley and Lang 1994; Svebak and Murgatroyd 1985), and enjoyment (Kendzierski and DeCarlo 1991; Tamborini et al. 2010). There are also other instruments not in the corpus that measure hedonic qualities of UX, such as the AttrakDiff questionnaire (Hassenzahl, Burmester, and Koller 2003) and the iScale survey tool (Karapanos et al. 2010). There were also a large number of audio properties studied, although the focus was generally on comparing different audio types or technologies. Future review papers might thus investigate these differences by focusing on different implementations of VR audio and/or different audio types.

The corpus also revealed four research design concerns that need to be addressed. The first is that of familiarisation tasks, since it was expected that the initial novelty that fades with continued use of VR technology might introduce a confounding factor in the results. While this is not specific to studies focusing on audio, the impact is expected to be more pronounced for audio studies since the technologies for providing audio feedback in VR are also relatively novel. However, the implications for future research are made more complex by the fact that VR use might also cause simulator sickness, i.e. another confounding factor. Guidelines for minimising these confounding factors would thus benefit future research in VR research and, by extension, VR audio research. The second research design concern pertains to the selection of stimuli, particularly musical stimuli. Music has the potential to affect aspects of experience for several reasons, such as preference, cultural connotations, rich and varied properties, etc., but there seems to be no standard protocol for selecting musical stimuli to be used in audio research. A starting point into such standardisation would be to conduct a systematic review focusing on the selection process for music towards a framework that might inform future research practices. Similarly, the effect of various display factors, such as the effect of different headphones or speakers, needs further investigation, which also calls for review papers focused on such technological differences.

The third research design concern is a lack of descriptive research that draws on knowledge outside laboratory environments. The markets for VR hardware and consumer software have already been a multi-million-dollar industry during the time during which the bulk of the corpus was published and has shown significant growth, which suggests an ample user base to draw from. Future research might thus use existing approaches to draw on sizable user bases, e.g. by deploying surveys via Reddit (Rogers and Weber 2019). Such an approach might not only improve the ecological validity of results, but also challenge the assumptions of researchers and practitioners whose audio-centric perspective might differ substantially from their target audience. Making use of existing VR content, such as games that are commercially available rather than creating software for each study can also save significant time and costs in conducting such research.

Finally, two more methodological concerns pertaining to the most-measured and second most-measured outcomes, i.e. presence and realism, were revealed. These concerns are not exclusive to the domain of audio but apply to VR research in general. Previous criticism has been leveled against measuring presence using purely psychometric self-reports with the suggestion toward combined approaches such as combining subjective self-reports and objective performance-based measures. The corpus reveals that this approach is underexplored, and future research should thus put further emphasis on such a combined approach. The list of analysis techniques used throughout the corpus

also shows a small number of approaches based on qualitative data, including thematic analysis and grounded theory but also singular instances of a linguistic inquiry word count, verbal protocol analysis, interaction analysis, cinemetrics, and analyzing transition diagrams. These and other techniques might provide fruitful alternatives for measuring presence that do not rely solely on subjective self-reports.

With regards to realism, the corpus reveals that the narrow view focusing on verisimilitude with real environments prevails over alternatives that allow for deviation from pure simulation. Only two measures that might be regarded as such alternatives were found in the corpus, i.e. naturalness and believability. Given the large market sector for VR applications in the entertainment sector, there is a greater need to understand the complex relationship between realism, unrealism, and willing suspension of disbelief in the context of VR audio. While audio implementations such as binaural and ambisonic audio are enjoying an increase in research and development, the perception of these technologies by the general public needs to be considered from a perspective that goes beyond recreating audio stimuli from physical reality (Rumsey 2014).

5.3. Practical implications

As the consumer market for VR applications continues to grow, there is much potential for exploring the boundaries of this still relatively new market. Some opportunities for audio-focused experiences highlighted by the corpus include musical listening, musical creation, mindfulness practice, and arts and cultural heritage. Previous studies exemplify the possibilities in terms of attending virtual concerts (Lind et al. 2017; Kasuya et al. 2019; Shin et al. 2019; Holm, Väänänen, and Battah 2020) as well as different approaches or tools for music creation in VR (Deacon, Stockman, and Barthet 2017; Chang, Kim, and Kim 2018). For these areas, the promise of co-located virtual experiences such as that of the Metaverse (Mystakidis 2022) also applies. For example, musical concerts are inherently social experiences and musical creation, including less traditionally "instrument-driven" genres such as techno and electronica, are often created collaboratively. Designers of VR applications can thus explore the possibilities of VR applications from a social-VR perspective.

Outside the area of audio-focused VR applications, there are several implications that can be asserted. Firstly, Thethe consistent findings regarding the effect of natural ambiences on certain experiential outcomes, such as pleastantness, calmness, and comfort also

highlight the potential of using VR to create mindfulness applications or other types of applications that aim to combat e.g., stress, and anxiety, an existing example of which is SoundSelf (https://www.soundself.com). The promise of co-located experiences might also extend into this domain, as mindfulness practice can often involve a guided aspect and active participation in music therapy provides unique affordances for different types of social-musical interactions. The use of VR also presents the potential to approach museum installations for arts and cultural heritage in entirely new ways, such as placing an audience "inside" a representation rather than mimicking a traditional exhibition space, the latter of which is still the prevalent approach reflected in the corpus. Finally, the enhanced gestural affordances of VR technology offers the potential to create versions of these and other applications that are audio-only or 'audio-mostly', which would be more accessible by those with visual impairments.

6. Limitations

For the sake of generalizability, we have adopted a wide scope, as reflected in the broad terms used for the search strings. The number of papers included as well as the variety of avenues and domains also indicates that the results were generally heterogeneous in nature, which provides a good indication of the generalizability of results. However, the scope was also purposefully limited to studies which made use of HMDs or audio-only solutions with head-tracking since it was reasoned that these technologies demand their own approach in terms of audio implementation. This excludes other types of immersive displays such as CAVEs from the current study. The search term also required studies to explicitly use the term 'experience' in the title, abstract, or keywords, although this excludes some studies that do not use this specific term but still studied some aspect related to user/player experience. No distinction was made between different audio types/roles, e.g. sound effects, music, etc., as the intention was to provide a broad overview of audio in this context. This does, however, limit insights that might be drawn from considering different types of audio separately. The search string used for ACM DL did not exclude papers based on the document type (as with the SCOPUS search string) and, as such, these were manually excluded during the screening process. The fact that only SCOPUS and ACM DL were used as databases might bias results against domain areas and disciplines that are underrepresented in these databases, such as game studies and humanities-based research. We also could not access all publication venues, which might

also bias results against certain domains, types of studies, etc. Lastly, while the PRISMA-Scr protocol was used to guide the process of searching, excluding, and reporting, we did not make use of an a-priori protocol for this process.

7. Conclusions

The effective use of audio in HMD-based VR presents a variety of advantages for inducing desired experiential outcomes. However, the effects of audio in various contexts are still quite poorly understood and research in this domain is marred by inconsistencies in measuring outcomes, selecting stimuli, and other research design concerns. This study presents a review of 121 studies that investigate the effect of audio on the user/player experience in HMD-based VR. The goal was to provide a broad overview of the domain areas where the experiential effects of audio are being studied, key methodological approaches for doing so, as well as how various outcomes are found to affect experience in HMD-based VR. The most common domains highlighted were that of entertainment, soundscape research, and therapy/rehabilitation as well as medium-specific studies on VR itself (answering RQ1: In which contexts are the effects of audio in HMD-based VR being studied?). Presence, realism, immersion, and pleasantness were found to be the most-commonly studied experiential aspects (answering RQ2: Which aspects of experience are being studied with regards to audio in HMD-based VR research?) with some of the most consistent outcomes; natural and artificial soundscapes are respectively perceived as pleasant and unpleasant, audio increases perceived realism but generally through a narrow lens of verisimilitude, and presence is largely unaffected by the introduction of audio in VR (answering RQ3: What insights does the literature research provide about the effect of audio on HMD-based VR experiences?). Directions for future research were drawn from these results, including the need for familiarisation tasks to reduce novelty bias and a validated questionnaire for measuring pleasantness. The current literature also reveals a dearth of research investigating social experiences as well as a lack of descriptive research drawing from everyday contexts, which might provide useful guidelines for designers of VR applications (answering RQ4: What implications from the literature can be drawn regarding the design of future HMD-based VR audio research?). Future work in the form of systematic reviews might analyze and discuss specific variables in more depth and consider different audio types or roles as well as qualitative data and implications these factors/insights

might have in the design of studies that investigate audio in HMD-based VR.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Isak de Villiers Bosman http://orcid.org/0000-0002-6840-9043
Oğuz 'Oz' Buruk http://orcid.org/0000-0002-8655-5327
Kristine Jørgensen http://orcid.org/0000-0002-6685-7598
Juho Hamari http://orcid.org/0000-0002-6573-588X

References

Aalmoes, R., and R. Zon. 2019. "Speech Intelligibility Test in Virtual Reality for Interior Noise Evaluation." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 259: 390–398.

Alzayat, A., M. Hancock, and M. A. Nacenta. 2019. "Quantitative Measurement of Tool Embodiment for Virtual Reality Input Alternatives." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–11. Association for Computing Machinery. https://doi.org/10.1145/3290605.3300673

Andreano, J., K. Liang, L. Kong, D. Hubbard, B. K. Wiederhold, and M. D. Wiederhold. 2009. "Auditory Cues Increase the Hippocampal Response to Unimodal Virtual Reality." *CyberPsychology and Behavior* 12 (3): 309–313. doi:10.1089/cpb.2009.0104.

Arksey, H., and L. Malley. 2005. "Scoping Studies: Towards a Methodological Framework." *International Journal of Social Research Methodology* 8 (1): 19–32. doi:10.1080/1364557032000119616.

Baka, E., M. Kentros, G. Papagiannakis, and N. Magnenat-Thalmann. 2018. "Virtual Reality Rehabilitation Based on Neurologic Music Therapy: A Qualitative Preliminary Clinical Study." In *Learning and Collaboration Technologies. Learning and Teaching*, edited by P. Zaphiris and A. Ioannou, 113–127. Springer International Publishing. doi:10.1007/978-3-319-91152-6_9.

Baker, H. 2022. "Meta: $1 Billion Spent In Quest Store." *UploadVR*. Accessed October 17, 2022. https://uploadvr.com/meta-1-billion-quest-store.

Bălan, O., M.- Dascălu, A. Moldoveanu, S. Caraiman, F. Ungureanu, R. Lupu, and A. Stan. 2018. "Brain Activation During Virtual and Real-World Mobility Tasks with Single and Multimodal Sensory Input Provided by an Assistive Navigational Device for Visually Impaired People." In *2018 25th Saint Petersburg*

International Conference on Integrated Navigation Systems (ICINS), 1–6. https://doi.org/10.23919/ICINS.2018.8405870.

Barlow, C., T. Michailidis, G. Meadow, J. Gouch, and E. Rajabally. 2019. "Using Immersive Audio and Vibration to Enhance Remote Diagnosis of Mechanical Failure in Uncrewed Vessels." Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio, York, UK. https://www.aes.org/e-lib/browse.cfm?elib=20428.

Bergstrom, I., S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater. 2017. "The Plausibility of a String Quartet Performance in Virtual Reality." *IEEE Transactions on Visualization and Computer Graphics* 23 (4): 1352–1359. https://doi.org/10.1109/TVCG.2017.2657138.

Bialkova, S., and M. S. V. Gisbergen. 2017. "When Sound Modulates Vision: VR Applications for Art and Entertainment." In *2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)*, 1–6. https://doi.org/10.1109/WEVR.2017.7957714.

Bird, J. M., C. I. Karageorghis, S. J. Baker, and D. A. Brookes. 2019. "Effects of Music, Video, and 360-Degree Video on Cycle Ergometer Exercise at the Ventilatory Threshold." *Scandinavian Journal of Medicine & Science in Sports* 29 (8): 1161–1173. https://doi.org/10.1111/sms.13453.

Bird, J. M., C. I. Karageorghis, S. J. Baker, D. A. Brookes, and A. V. Nowicky. 2020. "Ready Exerciser One: Effects of Music and Virtual Reality on Cycle Ergometer Exercise." *British Journal of Health Psychology* 26 (1): 15–32. https://doi.org/10.1111/bjhp.12445.

Blascovich, J., and J. Bailenson. 2011. *Infinite Reality: Avatars, Eternal Life, New Worlds, and the Dawn of the Virtual Revolution*. New York: William Morrow and Co.

Bombeke, K., A. Van Dongen, W. Durnez, A. Anzolin, H. Almgren, A. All, J. Van Looy, L. De Marez, D. Marinazzo, and E. P. Núñez Castellar. 2018. "Do Not Disturb: Psychophysiological Correlates of Boredom, Flow and Frustration During VR Gaming." In *Augmented Cognition: Intelligent Technologies*, edited by D. D. Schmorrow, and C. M. Fidopiastis, 101–119. Springer International Publishing. https://doi.org/10.1007/978-3-319-91470-1_10.

Bosman, I. de V., K. de Beer, and T. J. D. Bothma. 2021. "Creating Pseudo-Tactile Feedback in Virtual Reality Using Shared Crossmodal Properties of Audio and Tactile Feedback." *South African Computer Journal* 33 (1): Article 1. https://doi.org/10.18489/sacj.v33i1.883.

Bouvier, P. 2008. "The Five Pillars of Presence: Guidelines to Reach Presence." Proceedings of the 11th Annual International Workshop on Presence. Presence 2008, Padova. https://www.semanticscholar.org/paper/The-Five-Pillars-of-Presence%3A-Guidelines-to-Reach-Bouvier/979cf3f389c33e2748d1fe39122b1d653467911c

Bradley, M. M., and P. J. Lang. 1994. "Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential." *Journal of Behavior Therapy and Experimental Psychiatry* 25 (1): 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Brambilla, G., Di Gabriele. 2009. "Urban Soundscape Appraisal in Laboratory: A Tool for Testing and Improving Urban Renewal." *Proceedings of 8th European Conference on Noise Control*: 26–28.

Brinkman, W., A. Hoekstra, and R. van Egmond. 2015. "The Effect of 3D Audio and other Audio Techniques on Virtual Reality Experience." *Annual Review of Cybertherapy and Telemachine 2015, Advanced Technologies in the Behavioral Social and Neurosciences*, 44–48. doi:10.3233/978-1-61499-595-1-44.

Brinkman, W.-P., G. Sandino, and C. van der Mast. 2009. "Field Observations of Therapists Conducting Virtual Reality Exposure Treatment for the Fear of Flying." *European Conference on Cognitive Ergonomics: Designing beyond the Product --- Understanding Activity and User Experience in Ubiquitous Environments* 1–9.

Buele, J., V. M. López, J. Varela-Aldás, A. Soria, and G. Palacios-Navarro. 2020. "Virtual Environment Application that Complements the Treatment of Dyslexia (VEATD) in Children." In *Information Technology and Systems*, edited by Á. Rocha, C. Ferrás, C. E. Montenegro Marin and V. H. Medina García, 330–339. Springer International Publishing. doi:10.1007/978-3-030-40690-5_33.

Bujic, M., and J. Hamari. 2020. "Immersive Journalism: Extant Corpus and Future Agenda." *CCEUR Workshop Proc. CEUR Workshop Proceedings. Presented at the 4th International GamiFIN Conference* 136–145.

Byrns, A., H. Ben Abdessalem, M. Cuesta, M.-A. Bruneau, S. Belleville, and C. Frasson. 2020. "Adaptive Music Therapy for Alzheimer's Disease Using Virtual Reality." In *Intelligent Tutoring Systems*, edited by V. Kumar, and C. Troussas, 214–219. Springer International Publishing. https://doi.org/10.1007/978-3-030-49663-0_25.

Çamci, A. 2019. "Exploring the Effects of Diegetic and Non-diegetic Audiovisual Cues on Decision-making in Virtual Reality." /paper/Exploring-the-Effects-of-Diegetic-and-Non-diegetic-%C3%87amci/e3bf1083e493ba68b81d47e45f1d5295e17302d4.

Cavdir, D., and G. Wang. 2022. "Designing Felt Experiences with Movement-Based, Wearable Musical Instruments: From Inclusive Practices Toward Participatory Design." *Wearable Technologies* 3: e19. https://doi.org/10.1017/wtc.2022.15.

Chang, S. 2018. "Scoping Reviews and Systematic Reviews: Is It an Either/Or Question?." *Annals of Internal Medicine* 169 (7): 502–503. doi:10.7326/M18-2205.

Chang, M., Y. R. Kim, and G. J. Kim. 2018. "A Perceptual Evaluation of Generative Adversarial Network Real-Time Synthesized Drum Sounds in a Virtual Environment." In *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 144–148. https://doi.org/10.1109/AIVR.2018.00030.

Chen, H., A. Dey, M. Billinghurst, and R. W. Lindeman. 2017a. "Exploring Pupil Dilation in Emotional Virtual Reality Environments." In *Proceedings of the 27th International Conference on Artificial Reality and Telexistence and 22nd Eurographics Symposium on Virtual Environments*, 169–176.

Chen, H., A. Dey, M. Billinghurst, and R. W. Lindeman. 2017b. "Exploring the Design Space for Multi-Sensory Heart Rate Feedback in Immersive Virtual Reality." In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, 108–116. https://doi.org/10.1145/3152771.3152783.

Chen, T., Y.-S. Wu, and K. Zhu. 2018. "Investigating Different Modalities of Directional Cues for Multi-Task Visual-

Searching Scenario in Virtual Reality." In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 1–5. https://doi.org/10.1145/3281505.3281516.

Chirico, A., and A. Gaggioli. 2019. "Virtual-Reality Music-Based Elicitation of Awe: When Silence is Better than Thousands Sounds." In *Pervasive Computing Paradigms for Mental Health*, edited by P. Cipresso, S. Serino, and D. Villani, 1–11. Springer International Publishing. https://doi.org/10.1007/978-3-030-25872-6_1.

Chittaro, L. 2012. "Changing Heartbeat Perception to Induce Anxiety in Virtual Environments." *Studies in Health Technology and Informatics* 181: 172–176. doi:10.3233/978-1-61499-121-2-172.

Choi, J. H., and C. Diehm. 2021. "Aesthetic Flattening." *Interactions* 28 (4): 21–23. https://doi.org/10.1145/3468080.

Collins, K. 2013. *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games*. London: MIT Press.

Collins, K. 2020. "Game Sound: Reverberations." *Journal of Sound and Music in Games* 1 (1): 100–102. https://doi.org/10.1525/jsmg.2020.1.1.100.

Collins, K., and B. Kapralos. 2012. "Beyond the Screen: What We Can Learn About Game Design from Audio-Based Games." Proceedings of the Computer Games Multimedia and Allied Technology (CGAT 2012) conference, 5.

Cranmer, A., J. D. Ericson, A. Ebers Broughel, B. Bernard, E. Robicheaux, and M. Podolski. 2020. "Worth a Thousand Words: Presenting Wind Turbines in Virtual Reality Reveals New Opportunities for Social Acceptance and Visualization Research." *Energy Research & Social Science* 67: 101507. https://doi.org/10.1016/j.erss.2020.101507.

D'Ausilio, A., G. Novembre, L. Fadiga, and P. E. Keller. 2015. "What Can Music Tell us About Social Interaction?" *Trends in Cognitive Sciences* 19 (3): 111–114. https://doi.org/10.1016/j.tics.2015.01.005.

Davies, S., S. Cunningham, and R. Picking. 2017. "A Comparison of Audio Models for Virtual Reality Video." *2017 International Conference on Cyberworlds (CW)* 150–153. doi:10.1109/CW.2017.41.

Deacon, T., T. Stockman, and M. Barthet. 2017. "User Experience in an Interactive Music Virtual Reality System: An Exploratory Study." In *Bridging People and Sound*, edited by M. Aramaki, R. Kronland-Martinet, and S. Ystad, 192–216. Springer International Publishing. https://doi.org/10.1007/978-3-319-67738-5_12.

Desnoyers-Stewart, J., E. R. Stepanova, P. Pasquier, and B. E. Riecke. 2019. "JeL: Connecting Through Breath in Virtual Reality." In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3290607.3312845.

Dickinson, P., K. Gerling, K. Hicks, J. Murray, J. Shearer, and J. Greenwood. 2019. "Virtual Reality Crowd Simulation: Effects of Agent Density on User Experience and Behaviour." *Virtual Reality* 23 (1): 19–32. doi:10.1007/s10055-018-0365-0.

Difede, J., and H. G Hoffman. 2002. "Virtual Reality Exposure Therapy for World Trade Center Post-traumatic Stress Disorder: A Case Report." *CyberPsychology and Behavior* 5 (6): 529–535. doi:10.1089/109493102321018169.

Dong, M., H. Wang, and R. Guo. 2017. "Towards Understanding the Differences of Using 3D Auditory Feedback in Virtual Environments Between People with and Without Visual Impairments." In *2017 IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, 1–5. https://doi.org/10.1109/SIVE.2017.7901608.

Droste, M., J. Letellier, and J. Sieck. 2018. "An Interactive Classical VR Concert Featuring Multiple Views." In *Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities*, 1–4. https://doi.org/10.1145/3283458.3283501.

Dumlu, B. N., and Y. Demir. 2020. "Analyzing the User Experience of Virtual Reality Storytelling with Visual and Aural Stimuli." In *Design, User Experience, and Usability. Design for Contemporary Interactive Environments*, edited by A. Marcus and E. Rosenzweig, 409–425. Springer International Publishing. doi:10.1007/978-3-030-49760-6_29.

Eames, A. 2019. "Beyond Reality." In *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry, VRCAI '19*, 1–2. New York: Association for Computing Machinery. doi:10.1145/3359997.3365736.

Echevarria Sanchez, G. M., T. Van Renterghem, K. Sun, B. De Coensel, and D. Botteldooren. 2017a. "Using Virtual Reality for Assessing the Role of Noise in the Audio-Visual Design of an Urban Public Space." *Landscape and Urban Planning* 167: 98–107. https://doi.org/10.1016/j.landurbplan.2017.05.018.

Echevarria Sanchez, G. M., T. Van Renterghem, K. Sun, B. De Coensel, and D. Botteldooren. 2017b. "Personal Factors Affecting the Audio-Visual Perception of the Urban Public Space." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 255 (2): 5714–5723.

Ekman, I. 2005. "Meaningful Noise: Understanding Sound Effects in Computer Games." Proceedings of the 2005 Conference on Digital Arts & Culture, 5. https://www.researchgate.net/profile/Inger_Ekman/publication/224927551_Meaningful_Noise_Understanding_Sound_Effects_in_Computer_Games/links/546208640cf27487b4557a95.pdf.

Erfanian, M., A. J. Mitchell, J. Kang, and F. Aletta. 2019. "The Psychophysiological Implications of Soundscape: A Systematic Review of Empirical Literature and a Research Agenda." *International Journal of Environmental Research and Public Health* 16 (19): 3533. https://doi.org/10.3390/ijerph16193533.

Falconer, L. 2017. "Experiencing Sense of Place in Virtual and Physical Avebury." *Personal and Ubiquitous Computing* 21 (6): 977–988. https://doi.org/10.1007/s00779-017-1064-7.

Falconer, L., D. Burden, R. Cleal, R. Hoyte, P. Phelps, N. Slawson, N. Snashall, and K. Welham. 2020. "Virtual Avebury: Exploring Sense of Place in a Virtual Archaeology Simulation." *Virtual Archaeology Review* 11 (23): Article 23. doi:10.4995/var.2020.12924.

Farkas, T., S. Wiseman, P. Cairns, and R. Fiebrink. 2020. "A Grounded Analysis of Player-Described Board Game Immersion." In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 427–437. https://doi.org/10.1145/3410404.3414224.

Feng, M., A. Dey, and R. W. Lindeman. 2016. "An Initial Exploration of a Multi-Sensory Design Space: Tactile Support for Walking in Immersive Virtual Environments." In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, 95–104. https://doi.org/10.1109/3DUI.2016.7460037.

Fletcher, C., V. Hulusic, and P. Amelidis. 2019. "Virtual Reality Ear Training System: A Study on Spatialised Audio in Interval Recognition." In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 1–4. https://doi.org/10.1109/VS-Games.2019.8864592.

Gao, B., J.-I. Kim, and H. Kim. 2018. "Sensory and Perceptual Consistency for Believable Response in Action Feedback Loop." *Proceedings of Computer Graphics International* 2018: 201–210. https://doi.org/10.1145/3208159.3208171.

Garcia, D., M. Rychtáriková, and C. Glorieux. 2015. "Human Echolocation: Localizing Reflections of Self-generated Oral Sounds in Laboratory." *Proceedings of EuroNoise* 2015: 751–756.

Garner, T. A. 2017. *Echoes of Other Worlds: Sound in Virtual Reality: Past, Present and Future*. Cham, Switzerland: Springer.

Geronazzo, M., A. Rosenkvist, D. S. Eriksen, C. K. Markmann-Hansen, J. Køhlert, M. Valimaa, M. B. Vittrup, S. Serafin, and M. Picone. 2019. "Creating an Audio Story with Interactive Binaural Rendering in Virtual Reality." *Wireless Communications and Mobile Computing*. doi:10.1155/2019/1463204.

Geronazzo, M., E. Sikström, J. Kleimola, F. Avanzini, Götzen, A. de, and S. Serafin. 2018. "The Impact of an Accurate Vertical Localization with HRTFs on Short Explorations of Immersive Virtual Reality Scenarios." *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*: 90–97. doi:10.1109/ISMAR.2018.00034.

Geronazzo, M., S. Spagnol. 2013. "Mixed Structural Modeling of Head-related Transfer Functions for Customized Binaural Audio Delivery." *18th International Conference on Digital Signal Processing (DSP)* 1–8. doi:10.1109/ICDSP.2013.6622764.

Ghosh, S., 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): 1447–1456. https://doi.org/10.1109/TVCG.2018.2793698.

Gilbert, S. B. 2016. "Perceived Realism of Virtual Environments Depends on Authenticity." *Presence: Teleoperators and Virtual Environments* 25 (4): 322–324. https://doi.org/10.1162/PRES_a_00276.

Gomez, J., H. G. Hoffman, S. L. Bistricky, M. Gonzalez, L. Rosenberg, M. Sampaio, A. Garcia-Palacios, et al. 2017. "The Use of Virtual Reality Facilitates Dialectical Behavior Therapy® "Observing Sounds and Visuals" Mindfulness Skills Training Exercises for a Latino Patient with Severe Burns: A Case Study." *Frontiers in Psychology* 8. https://doi.org/10.3389/fpsyg.2017.01611.

Gospodarek, M., A. Genovese, D. Dembeck, C. Brenner, A. Roginska, and K. Perlin. 2019. "Sound Design and Reproduction Techniques for Co-Located Narrative VR Experiences." *Audio Engineering Society Convention* 147. https://www.aes.org/e-lib/browse.cfm?elib=20660.

Grand View Research. 2020. "Virtual Reality in Gaming Market Size Report, 2020–2027." https://www.grandviewresearch.com/industry-analysis/virtual-reality-in-gaming-market.

Guang-yu, Z., Z. Yu-an, and G. Haochen. 2019. "Experimental Research on Influencing Factors of Urban Street Traffic Acoustic Annoyance Based on VR Experience." In *26th International Congress on Sound and Vibration 2019 (ICSV26)*, 1582–1589. https://www.iiav.org/content/archives_icsv_last/2019_icsv26/content/papers/papers/full_paper_252_20190524212555907.pdf.

Guillen, G., H. Jylhä, and L. Hassan. 2021. "The Role Sound Plays in Games: A Thematic Literature Study on Immersion, Inclusivity and Accessibility in Game Sound Research." *Academic Mindtrek* 2021: 12–20. doi:10.1145/3464327.3464365.

Harmon-Jones, C., B. Bastian, and E. Harmon-Jones. 2016. "The Discrete Emotions Questionnaire: A New Tool for Measuring State Self-Reported Emotions." *PLOS ONE* 11 (8): e0159915. https://doi.org/10.1371/journal.pone.0159915.

Harriet, S., and D. T. Murphy. 2015. "Auralisation of an Urban Soundscape." *Acta Acustica United with Acustica* 101 (4): 798–810. https://doi.org/10.3813/AAA.918874.

Hassenzahl, M., M. Burmester, and F. Koller. 2003. "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität." In *Mensch & Computer 2003: Interaktion in Bewegung*, edited by G. Szwillus, and J. Ziegler, 187–196. Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-80058-9_19.

Heineken, E., and F. P. Schulte. 2007. "Seeing Size and Feeling Weight: The Size-Weight Illusion in Natural and Virtual Reality." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49 (1): 136–144. https://doi.org/10.1518/001872007779598028.

Hirzle, T., M. Cordts, E. Rukzio, J. Gugenheimer, and A. Bulling. 2021. "A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445361.

Høeg, E. R., T. M. Povlsen, J. R. Bruun-Pedersen, B. Lange, N. C. Nilsson, K. B. Haugaard, S. M. Faber, S. W. Hansen, C. K. Kimby, and S. Serafin. 2021. "System Immersion in Virtual Reality- Based Rehabilitation of Motor Function in Older Adults: A Systematic Review and Meta-Analysis." *Frontiers in Virtual Reality*. doi:10.3389/frvir.2021.647993.

Holm, J., K. Väänänen, and A. Battah. 2020. "User Experience of Stereo and Spatial Audio in 360°; Live Music Videos." In *Proceedings of the 23rd International Conference on Academic Mindtrek*, 134–141. https://doi.org/10.1145/3377290.3377291.

Hong, D., H.-J. Kwon, C. G. Kim, and W.-C. Park. 2018. "Real-time 3D Audio Downmixing System based on Sound Rendering for the Immersive Sound of Mobile Virtual Reality Applications." *KSII Transactions on Internet and Information Systems : TIIS* 12 (12): 5936–5954.

Hong, J.-Y., B. Lam, Z.-T. Ong, W.-S. Gan, J. Fen, and S.-T. Tan. 2019. "Effect of Spatial Separation Between Traffic Noise and Water Sound on Soundscape Assessment." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 259: 5250–5255. doi:10.3837/tiis.2018.12.018.

Hoppe, M., B. Rossmy, D. P. Neumann, S. Streuber, A. Schmidt, and T.-K. Machulla. 2020. "A Human Touch: Social Touch Increases the Perceived Human-Likeness of

Agents in Virtual Reality." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–11. https://doi.org/10.1145/3313831.3376719.

Hruby, F. 2019. "The Sound of Being There: Audiovisual Cartography with Immersive Virtual Environments." *KN - Journal of Cartography and Geographic Information* 69 (1): 19–28. doi:10.1007/s42489-019-00003-5.

Huang, H., M. Solah, D. Li, and L.-F. Yu. 2019. "Audible Panorama: Automatic Spatial Audio Generation for Panorama Imagery." *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* 1–11. doi:10.1145/3290605.3300851.

Ibrahim, E. N. M., and W. N. Wan Bashir. 2019. "Enhancing Food Taste Experience with 360° Immersive Virtual Reality (IVR)." *International Journal of Advanced Trends in Computer Science and Engineering* 8 (1.6 Sp): Article 1.6 Sp. doi:10.30534/ijatcse/2019/5781.62019.

IJsselsteijn, W. A., Y. A. W. de Kort, K. Poels, and Human Technology Interaction. 2013. "The Game Experience Questionnaire." *Technische Universiteit Eindhoven*. https://research.tue.nl/nl/publications/3d28afa4-d1c8-412d-be83-654f00dcae0f

Ijsselsteijn, W. A., H. de Ridder, J. Freeman, and S. E. Avons. 2000. "Presence: Concept, Determinants, and Measurement." *Human Vision and Electronic Imaging V* 3959: 520–529. https://doi.org/10.1117/12.387188.

Isbister, K. 2016. *How Games Move Us: Emotion by Design.* London: MIT Press.

Ishikawa, A. 2019. "Visual and Auditory Impression for Virtual Reality Space Expressed by Panoramic Image and Impulse Response Signal." *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)* 1084–1087. doi:10.1109/GCCE46687.2019.9015489.

Jain, D., S. Junuzovic, E. Ofek, M. Sinclair, J. Porter, C. Yoon, S. Machanavajhala, and M. Ringel Morris. 2021. "A Taxonomy of Sounds in Virtual Reality." *Designing Interactive Systems Conference* 2021: 160–170. doi:10.1145/3461778.3462106.

Jain, D., M. Sra, J. Guo, R. Marques, R. Wu, J. Chiu, and C. Schmandt. 2016. "Immersive Scuba Diving Simulator Using Virtual Reality." *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* 729–739. doi:10.1145/2984511.2984519.

Janer, J., E. Gomez, A. Martorell, M. Miron, and B. de Wit. 2016. "Immersive Orchestras: Audio Processing for Orchestral Music VR Content." *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, 1–2. Barcelona, Spain: Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/VS-GAMES.2016.7590352.

Jenny, C., and C. Reuter. 2021. "Can I Trust My Ears in VR? Head-related Transfer Functions and Valuation Methods with Descriptive Attributes in Virtual Reality." *International Journal of Virtual Reality* 21 (2): Article 2. doi:10.20870/IJVR.2021.21.2.4831.

Jeon, J. Y., and H. I. Jo. 2019. "Three-dimensional Virtual Reality-Based Subjective Evaluation of Road Traffic Noise Heard in Urban High-Rise Residential Buildings." *Building and Environment* 148: 468–477. https://doi.org/10.1016/j.buildenv.2018.11.004.

Jo, H. I., and J. Y. Jeon. 2020. "The Influence of Human Behavioral Characteristics on Soundscape Perception in Urban Parks: Subjective and Observational Approaches." *Landscape and Urban Planning* 203: 103890. https://doi.org/10.1016/j.landurbplan.2020.103890.

Johnston, D., H. Egermann, and G. Kearney. 2020. "SoundFields: A Virtual Reality Game Designed to Address Auditory Hypersensitivity in Individuals with Autism Spectrum Disorder." *Applied Sciences* 10 (9): 2996. https://doi.org/10.3390/app10092996.

Jørgensen, K. 2006. "On the Functional Aspects of Computer Game Audio." *Proceedings of the Audio Mostly Conference.* Audio Mostly Conference, Interactive Institute, Piteå, Sweden. https://bora.uib.no/bora-xmlui/handle/1956/6734.

Jørgensen, K. 2011. "Time for New Terminology?: Diegetic and Non-Diegetic Sounds in Computer Games Revisited." *Game Sound Technology and Player Interaction: Concepts and Developments* 78–97. doi:10.4018/978-1-61692-828-5.ch005.

Jørgensen, K. 2017. "Left in the Dark: Playing Computer Games with the Sound Turned Off." In *From Pac-Man to Pop Music*. 1st ed., edited by K. Collins, 163–176. Routledge.

Josman, N., A. Reisberg, P. L. Weiss, A. Garcia-Palacios, and H. G. Hoffman. 2008. "BusWorld: An Analog Pilot Test of a Virtual Environment Designed to Treat Posttraumatic Stress Disorder Originating from a Terrorist Suicide Bomb Attack." *CyberPsychology & Behavior* 11 (6): 775–777. https://doi.org/10.1089/cpb.2008.0048.

Kapralos, B., K. Collins, and A. Uribe-Quevedo. 2017. "The Senses and Virtual Environments." *The Senses and Society* 12 (1): 69–75. doi:10.1080/17458927.2017.1268822.

Karapanos, E., J. Zimmerman, J. Forlizzi, and J.-B. Martens. 2010. "Measuring the Dynamics of Remembered Experience Over Time." *Interacting with Computers* 22 (5): 328–335. https://doi.org/10.1016/j.intcom.2010.04.003.

Kasuya, T., M. Tsukada, Y. Komohara, S. Takasaka, T. Mizuno, Y. Nomura, Y. Ueda, and H. Esaki. 2019. "LiVRation: Remote VR Live Platform with Interactive 3D Audio-Visual Service." In *2019 IEEE Games, Entertainment, Media Conference (GEM)*, 1–7. https://doi.org/10.1109/GEM.2019.8811549.

Kauhanen, O., H. Väätäjä, M. Turunen, T. Keskinen, E. Sirkkunen, T. Uskali, V. Lindqvist, C. Kelling, and J. Karhu. 2017. "Assisting Immersive Virtual Reality Development with User Experience Design Approach." In *Proceedings of the 21st International Academic Mindtrek Conference*, 127–136. https://doi.org/10.1145/3131085.3131126.

Kelling, C., O. Kauhanen, H. Väätäjä, J. Karhu, M. Turunen, and V. Lindqvist. 2018. "Implications of Audio and Narration in the User Experience Design of Virtual Reality." In *Proceedings of the 22nd International Academic Mindtrek Conference*, 258–261. https://doi.org/10.1145/3275116.3275153.

Kendzierski, D., and K. J. DeCarlo. 1991. "Physical Activity Enjoyment Scale: Two Validation Studies." *Journal of Sport and Exercise Psychology* 13 (1): 50–64. https://doi.org/10.1123/jsep.13.1.50

Kern, A. C., and W. Ellermeier. 2018. "Influence of Hearing Your Steps and Environmental Sounds in VR While Walking." *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, 1–4. doi:10.1109/SIVE.2018.8577177.

Kern, A., and W. Ellermeier. 2020. "Audio in VR: Effects of a Soundscape and Movement-Triggered Step Sounds on Presence." *Frontiers in Robotics and AI* 7. https://doi.org/10.3389/frobt.2020.00020.

Kern, A. C., W. Ellermeier, and L. Jost. 2020. "The Influence of Mood Induction by Music or a Soundscape on Presence and Emotions in a Virtual Reality Park Scenario." *Proceedings of the 15th International Conference on Audio Mostly* 233–236. doi:10.1145/3411109.3411129.

Kim, Y., and H. Lee. 2022. "Falling in Love with Virtual Reality Art: A New Perspective on 3D Immersive Virtual Reality for Future Sustaining Art Consumption." *International Journal of Human–Computer Interaction* 38 (4): 371–382. https://doi.org/10.1080/10447318.2021.1944534.

Kim, Y. M., I. Rhiu, and M. H. Yun. 2020. "A Systematic Review of a Virtual Reality System from the Perspective of User Experience." *International Journal of Human–Computer Interaction* 36: 893–910. https://doi.org/10.1080/10447318.2019.1699746.

Kiss, F., S. Mayer, and V. Schwind. 2020. "Audio VR: Did Video Kill the Radio Star?." *Interactions* 27: 46–51. doi:10.1145/3386385.

Kong, Y., C. Sharma, M. Kanala, M. Thakur, L. Li, D. Xu, R. Harrison, and D. D. Torrico. 2020. "Virtual Reality and Immersive Environments on Sensory Perception of Chocolate Products: A Preliminary Study." *Foods (basel, Switzerland)* 9 (4): 515. https://doi.org/10.3390/foods9040515.

Kramer, A. D. I., L. M. Oh, and S. R. Fussell. 2006. "Using Linguistic Features to Measure Presence in Computer-Mediated Communication." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 913–916. https://doi.org/10.1145/1124772.1124907.

Kruijff, E., A. Marquardt, C. Trepkowski, R. W. Lindeman, A. Hinkenjann, J. Maiero, and B. E. Riecke. 2016. "On Your Feet! Enhancing Vection in Leaning-Based Interfaces Through Multisensory Stimuli." In *Proceedings of the 2016 Symposium on Spatial User Interaction*, 149–158. https://doi.org/10.1145/2983310.2985759.

Kugler, L. 2021. "The State of Virtual Reality Hardware." *Communications of the ACM* 64 (2): 15–16. https://doi.org/10.1145/3441290.

Kurabayashi, H., M. Otani, K. Itoh, M. Hashimoto, and M. Kayama. 2014. "Sound Image Localization Using Dynamic Transaural Reproduction with Non-contact Head Tracking." *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences* E97-A (9): 1849–1858.

Larsson, P., A. Väljamäe, D. Västfjäll, A. Tajadura-Jiménez, and M. Kleiner. 2010. "Auditory-Induced Presence in Mixed Reality Environments and Related Technology." In *The Engineering of Mixed Reality Systems*, edited by E. Dubois, P. Gray, and L. Nigay, 143–163. Springer London. https://doi.org/10.1007/978-1-84882-733-2_8.

Latoschik, M. E., F. Kern, J.-P. Stauffert, A. Bartl, M. Botsch, and J.-L Lugrin. 2019. "Not Alone Here?! Scalability and User Experience of Embodied Ambient Crowds in Distributed Social Virtual Reality." *IEEE Transactions on Visualization and Computer Graphics* 25 (5): 2134–2144. doi:10.1109/TVCG.2019.2899250.

Law, E. L.-C., F. Brühlmann, and E. D. Mekler. 2018. "Systematic Review and Validation of the Game Experience Questionnaire (GEQ)—Implications for Citation and Reporting Practice." In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, 257–270. https://doi.org/10.1145/3242671.3242683.

Lee, M., G. Bruder, and G. F. Welch. 2017. "Exploring the Effect of Vibrotactile Feedback Through the Floor on Social Presence in an Immersive Virtual Environment." In *2017 IEEE Virtual Reality (VR)*, 105–111. https://doi.org/10.1109/VR.2017.7892237.

Lee, C., H. Joo, and S. Jun. 2021. "Social VR as the New Normal? Understanding User Interactions for the Business Arena." In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–5. Association for Computing Machinery. https://doi.org/10.1145/3411763.3451740.

Lee, S., and J.-H. Lee. 2017. "Does Spatial Attribute Between 2D and 3D Virtual Spaces Make Different User Immersion of Audio-Visual Events?" In *Proceedings of the 9th International Conference on Computer and Automation Engineering*, 56–59. https://doi.org/10.1145/3057039.3057092.

Letwory, N., R. Aalmoes, and M. Miltenburg. 2018. "Perception and Presence in Virtual Reality for Simulated Aircraft Noise." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 258: 1716–1725.

Li, Y., P. Tennent, and S. Cobb. 2019. "Appropriate Control Methods for Mobile Virtual Exhibitions." In *VR Technologies in Cultural Heritage*, edited by M. Duguleană, M. Carrozzino, M. Gams, and I. Tanea, 165–183. Springer International Publishing. https://doi.org/10.1007/978-3-030-05819-7_13.

Li, H., and H. Xie. 2018. "Physiological Effects of Soundscape Perception in a Mountainous Urban Park: Comparison between in situ and Laboratory Methodologies." *25th International Congress on Sound and Vibration 2018 (ICSV 25): Hiroshima Calling*: 1603–1608.

Liao, H., N. Xie, H. Li, Y. Li, J. Su, F. Jiang, W. Huang, and H. T. Shen. 2020. "Data-Driven Spatio-Temporal Analysis via Multi-Modal Zeitgebers and Cognitive Load in VR." In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 473–482. https://doi.org/10.1109/VR46266.2020.00068.

Liljedahl, M., N. Papworth, and S. Lindberg. 2007. "Beowulf: A Game Experience Built on Sound Effects." https://smartech.gatech.edu/handle/1853/49983.

Lim, T., S. Lee, and F. Ke. 2017. "Integrating Music into Math in a Virtual Reality Game: Learning Fractions." *International Journal of Game-Based Learning (IJGBL)* 7 (1): 57–73. doi:10.4018/IJGBL.2017010104.

Lind, R. B., V. Milesen, D. M. Smed, S. P. Vinkel, F. Grani, N. C. Nilsson, L. Reng, R. Nordahl, and S. Serafin. 2017. "Sound Design in Virtual Reality Concert Experiences Using a Wave Field Synthesis Approach." In *2017 IEEE Virtual Reality (VR)*, 363–364. https://doi.org/10.1109/VR.2017.7892327.

Lindquist, M., B. Maxim, J. Proctor, and F. Dolins. 2020. "The Effect of Audio Fidelity and Virtual Reality on the Perception of Virtual Greenspace." *Landscape and Urban Planning* 202: 103884. doi:10.1016/j.landurbplan.2020.103884.

Maffei, L., M. Masullo, F. Aletta, and M. Di Gabriele. 2013. "The Influence of Visual Characteristics of Barriers on

Railway Noise Perception." *Science of The Total Environment* 445-446: 41–47. https://doi.org/10.1016/j.scitotenv.2012.12.025.

Mäkinen, H., E. Haavisto, S. Havola, and J.-M. Koivisto. 2022. "User Experiences of Virtual Reality Technologies for Healthcare in Learning: An Integrative Review." *Behaviour & Information Technology* 41: 1–17. https://doi.org/10.1080/0144929X.2020.1788162.

Masullo, M., L. Maffei, and R. Pellegrino. 2019. "An Experimental Study on the Audio-Visual Impact and Integration of Split System Units on Buildings Faades." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 259 (6): 3351–3357.

McArthur, A. 2016. "Disparity in Horizontal Correspondence of Sound and Source Positioning: The Impact on Spatial Presence for Cinematic VR." Audio engineering society conference: 2016 AES international conference on audio for virtual and augmented reality, September 21. https://www.aes.org/e-lib/browse.cfm?elib=18500.

Mcarthur, A., R. Stewart, and M. Sandler. 2017. "Sounds Too True to be Good: Diegetic Infidelity - The Case for Sound in Virtual Reality." *Journal of Media Practice* 18: 26–40. doi:10.1080/14682753.2017.1305840.

Mehra, R., A. Rungta, A. Golas, M. Lin, and D. Manocha. 2015. "WAVE: Interactive Wave-based Sound Propagation for Virtual Environments." *IEEE Transactions on Visualization and Computer Graphics* 21 (4): 434–442. doi:10.1109/TVCG.2015.2391858.

Men, L., N. Bryan-Kinns, A. S. Hassard, and Z. Ma. 2017. "The Impact of Transitions on User Experience in Virtual Reality." *2017 IEEE Virtual Reality (VR),* 285–286. doi:10.1109/VR.2017.7892288.

Michailidis, T., C. Barlow, G. Meadow, J. Gouch, and E. Rajabally. 2019. "Enhance Engine Room Diagnostics Through Audio-Focused VR Simulation." In *Human-Computer Interaction - INTERACT 2019*, edited by D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler and P. Zaphiris, 386–394. Springer International Publishing. doi:10.1007/978-3-030-29390-1_21.

Moraes, A. N., R. Flynn, A. Hines, and N. Murray. 2020. "Evaluating the User in a Sound Localisation Task in a Virtual Reality Application." In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 1–6. https://doi.org/10.1109/QoMEX48832.2020.9123136.

Munn, Z., M. D. J. Peters, C. Stern, C. Tufanaru, A. Mcarthur. 2018. "Systematic Review or Scoping Review? Guidance for Authors when Choosing between a Systematic or Scoping Review Approach." *BMC Medical Research Methodology* 18 (1): 143. doi:10.1186/s12874-018-0611-x.

Mystakidis, S. 2022. "Metaverse." *Encyclopedia* 2 (1): 486–497. https://doi.org/10.3390/encyclopedia2010031.

Nararro-Haro, M. V., H. G. Hoffman, A. Garcia-Palacios, M. Sampaio, W. Alhalabi, K. Hall, and M. Linehan. 2016. "The Use of Virtual Reality to Facilitate Mindfulness Skills Training in Dialectical Behavioral Therapy for Borderline Personality Disorder: A Case Study." *Frontiers in Psychology* 7. https://doi.org/10.3389/fpsyg.2016.01573.

Narciso, D., M. Bessa, J. Vasconcelos-Raposo, D. Narciso, M. Bessa, M. Melo, A. Coelho, J. Vasconcelos-Raposo, and A. Coelho. 2019. "Immersive 360° Video User Experience: Impact of Different Variables in the Sense of Presence and

Cybersickness." *Universal Access in the Information Society* 18 (1): 77–87. https://doi.org/10.1007/s10209-017-0581-5

Noisternig, M., T. Musil, A. Sontacchi, and R. Holdrich. 2003. "3D Binaural Sound Reproduction using a Virtual Ambisonic Approach." *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS '03* 174–178. doi:10.1109/VECIMS.2003.1227050.

Nordahl, R., and N. C. Nilsson. 2014. "The Sound of Being There." In *Oxford Handbook of Interactive Audio*. https://doi.org/10.1093/oxfordhb/9780199797226.013.013.

North, M., J. Hill, A. Aikhuele, and S. North. 2008. "Virtual Reality Training in Aid of Communication Apprehension in Classroom Environments." *International Journal of Emerging Technologies in Learning (IJET)* 3 (2): 34–37. doi:10.3991/ijet.v3i2.163.

Oberman, T., B. Bojanić Obad Šćitaroci, and K. Jambrošić. 2018. "Towards a Virtual Soundwalk." In *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*, 317–343. Hershey, PA: IGI Global. doi:10.4018/978-1-5225-3637-6.ch014.

Oh, C., F. Herrera, and J. Bailenson. 2019. "The Effects of Immersion and Real-World Distractions on Virtual Social Interactions." *Cyberpsychology, Behavior, and Social Networking* 22 (6): 365–372. https://doi.org/10.1089/cyber.2018.0404.

O'Hagan, J., J. R. Williamson, and M. Khamis. 2020. "Bystander Interruption of VR Users." In *Proceedings of the 9TH ACM International Symposium on Pervasive Displays*, 19–27. https://doi.org/10.1145/3393712.3395339.

Olko, M., D. Dembeck, Y.-H. Wu, A. Genovese, and A. Roginska. 2017. "Identification of Perceived Sound Quality Attributes of 360° Audiovisual Recordings in VR Using a Free Verbalization Method." Audio Engineering Society Convention 143, October 8. https://www.aes.org/e-lib/browse.cfm?elib=19227.

Ong, Z. T., J. Y. Hong, B. Lam, R. Gupta, W. S. Gan, S. H. Chong, J. Feng, and K. Ooi. 2017. "Effect of Masker Orientation on Masking Efficacy in Soundscape Applications." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 255: 4916–4922.

Ooi, K., J.-Y. Hong, B. Lam, Z. T. Ong, and W.-S. Gan. 2017. "Validation of Binaural Recordings with Head Tracking for Use in Soundscape Evaluation." *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 255: 4651–4659. https://dr.ntu.edu.sg//handle/10356/88176.

Optale, G., S. Capodieci, P. Pinelli, D. Zara, L. Gamberini, and G. Riva. 2001. "Music-enhanced Immersive Virtual Reality in the Rehabilitation of Memory related Cognitive Processes and Functional Abilities: A Case Report." *Presence: Teleoperators and Virtual Environments* 10 (4): 450–462. doi:10.1162/1054746011470217.

Palani, H., and N. A. Giudice. 2014. "Evaluation of Non-visual Panning Operations using Touch-screen Devices." *Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*: 293–294. doi:10.1145/2661334.2661336.

Papadopoulos, K., M. Barouti, and E. Koustriava. 2016. "The Improvement of Cognitive Maps of Individuals with Blindness Through the Use of an Audio-tactile Map." In *Universal Access in Human-Computer Interaction. Interaction Techniques and Environments*, edited by M.

Antona and C. Stephanidis, 72–80. Springer International Publishing. doi:https://doi.org/10.1007/978-3-319-40244-4_8

Park, S. H., P. J. Lee, T. Jung, and A. Swenson. 2020. "Effects of the Aural and Visual Experience on Psycho-Physiological Recovery in Urban and Rural Environments." *Applied Acoustics* 169: 107486. https://doi.org/10.1016/j.apacoust.2020.107486.

Parker, E., and M. Saker. 2020. "Art Museums and the Incorporation of Virtual Reality: Examining the Impact of VR on Spatial and Social Norms." *Convergence: The International Journal of Research Into New Media Technologies* 26 (5–6): 1159–1173. https://doi.org/10.1177/1354856519897251.

Pater, J., A. Coupe, R. Pfafman, C. Phelan, T. Toscos, and M. Jacobs. 2021. "Standardizing Reporting of Participant Compensation in HCI: A Systematic Literature Review and Recommendations for the Field." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–16. Association for Computing Machinery.

Pedersen, K., V. Hulusic, P. Amelidis, and T. Slattery. 2020. "Spatialized Audio in a Custom-Built OpenGL-Based Ear Training Virtual Environment." *IEEE Computer Graphics and Applications* 40 (5): 67–81. https://doi.org/10.1109/MCG.2020.3013439.

Peng, Y.-H., C. Yu, S.-H. Liu, C.-W. Wang, P. Taele, N.-H. Yu, and M. Y. Chen. 2020. "WalkingVibe: Reducing Virtual Reality Sickness and Improving Realism While Walking in VR Using Unobtrusive Head-Mounted Vibrotactile Feedback." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376847.

Phillips, W. 2018. "Composing Video Game Music for Virtual Reality: Diegetic versus Non-diegetic." June 14 https://www.gamasutra.com/blogs/WinifredPhillips/20180614/319894/Composing_video_game_music_for_Virtual_Reality_Diegetic_versus_Nondiegetic.php.

Picinali, L., A. Afonso, M. Denis, and B. F. G. Katz. 2014. "Exploration of Architectural Spaces by Blind People Using Auditory Virtual Reality for the Construction of Spatial Knowledge." *International Journal of Human-Computer Studies* 72 (4): 393–407. https://doi.org/10.1016/j.ijhcs.2013.12.008.

Pizzoli, S. F. M., S. Triberti, D. Monzani, K. Mazzocco, E. Kufel, M. Porebiak, and G. Pravettoni. 2019. "Comparison of Relaxation Techniques in Virtual Reality for Breast Cancer Patients." In *2019 5th Experiment International Conference (Exp.at'19)*, 348–351. https://doi.org/10.1109/EXPAT.2019.8876542.

Poudyal, A., A. van Heerden, A. Hagaman, C. Islam, A. Thapa, S. M. Maharjan, P. Byanjankar, and B. A. Kohrt. 2021. "What Does Social Support Sound Like? Challenges and Opportunities for Using Passive Episodic Audio Collection to Assess the Social Environment." *Frontiers in Public Health* 9: 633606. https://doi.org/10.3389/fpubh.2021.633606.

Prpa, M., K. Tatar, J. Françoise, B. Riecke, T. Schiphorst, and P. Pasquier. 2018. "Attending to Breath: Exploring How the Cues in a Virtual Environment Guide the Attention to Breath and Shape the Quality of Experience to Support Mindfulness." In *Proceedings of the 2018 Designing Interactive Systems Conference*, 71–84. https://doi.org/10.1145/3196709.3196765.

Prpa, M., K. Tatar, B. E. Riecke, and P. Pasquier. 2017. "The Pulse Breath Water System: Exploring Breathing as an Embodied Interaction for Enhancing the Affective Potential of Virtual Reality." In *Virtual, Augmented and Mixed Reality*, edited by S. Lackey, and J. Chen, 153–172. Springer International Publishing. https://doi.org/10.1007/978-3-319-57987-0_13.

Radianti, J., T. A. Majchrzak, J. Fromm, and I. Wohlgenannt. 2020. "A Systematic Review of Immersive Virtual Reality Applications for Higher Education: Design Elements, Lessons Learned, and Research Agenda." *Computers & Education* 147: 103778. https://doi.org/10.1016/j.compedu.2019.103778.

Rahman, J. S., T. Gedeon, S. Caldwell, and R. L. Jones. 2021. "Can Binaural Beats Increase Your Focus? Exploring the Effects of Music in Participants' Conscious and Brain Activity Responses." In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–6. Association for Computing Machinery. https://doi.org/10.1145/3411763.3451813.

Rees-Jones, J., and H. Daffern. 2019. "The Hills are Alive: Capturing and Presenting an Outdoor Choral Performance for Virtual Reality." Audio engineering society conference: 2019 AES international conference on immersive and interactive audio, March 17. https://www.aes.org/e-lib/browse.cfm?elib=20447

Ribeiro, G., K. Rogers, M. Altmeyer, T. Terkildsen, and L. E. Nacke. 2020. "Game Atmosphere: Effects of Audiovisual Thematic Cohesion on Player Experience and Psychophysiology." *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 107–119. https://doi.org/10.1145/3410404.3414245.

Robotham, T., O. Rummukainen, M. Kurz, M. Eckert, and E. Habets. 2022. "Comparing Direct and Indirect Methods of Audio Quality Evaluation in Virtual Reality Scenes of Varying Complexity." *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/TVCG.2022.3150491.

Rocchesso, D., S. Delle Monache, and S. Barrass. 2019. "Interaction by Ear." *International Journal of Human-Computer Studies* 131: 152–159. doi:10.1016/j.ijhcs.2019.05.012.

Rogers, K. 2017. "Exploring the Role of Audio in Games." *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '17 Extended Abstracts*: 727–731. doi:10.1145/3130859.3133227.

Rogers, K., S. Karaosmanoglu, M. Altmeyer, A. Suarez, and L. E. Nacke. 2022. "Much Realistic, Such Wow! A Systematic Literature Review of Realism in Digital Games." In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, 1–21. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3491102.3501875.

Rogers, K., M. Milo, M. Weber, and L. E. Nacke. 2020. "The Potential Disconnect between Time Perception and Immersion: Effects of Music on VR Player Experience." *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*: 414–426. doi:10.1145/3410404.3414246.

Rogers, K., and L. Nacke. 2017. "Exploring the Potential of Game Audio for Wellbeing." In *Proceedings of the Positive Gaming: Workshop on Gamification and Games for Wellbeing Co-located with the ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (CHI PLAY 2017)*, 1–7. Amsterdam, The Netherlands.

Rogers, K., G. Ribeiro, R. R. Wehbe, M. Weber, and L. E. Nacke. 2018. "Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3173902.

Rogers, K., and M. Weber. 2019. "Audio Habits and Motivations in Video Game Players." In *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*, 45–52. https://doi.org/10.1145/3356590.3356599.

Roohi, S., J. Takatalo, J. M. Kivikangas, and P. Hämäläinen. 2018. "Neural Network Based Facial Expression Analysis of GameEvents: A Cautionary Tale." In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, 429–437. https://doi.org/10.1145/3242671.3242701.

Rossi, H. S., S. M. Santos, R. O. Prates, and R. A. C. Ferreira. 2018. "Imaginator: A Virtual Reality Based Game for the Treatment of Sensory Processing Disorders." *IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)* 1–8. doi:10.1109/SeGAH.2018.8401355.

Rummukainen, O., S. Schlecht, A. Plinge, and E. A. P. Habets. 2017. "Evaluation of Binaural Reproduction Systems from Behavioral Patterns in a Six-degrees-of-freedom Wayfinding Task." *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)* 1–3. doi:10.1109/QoMEX.2017.7965680.

Rumsey, F. 2014. "Spatial Audio—Reconstructing Reality or Creating Illusion?" May 21. https://www.youtube.com/watch?v=y82nth2Pnwk.

Rzayev, R., G. Karaman, K. Wolf, N. Henze, and V. Schwind. 2019. "The Effect of Presence and Appearance of Guides in Virtual Reality Exhibitions." *Proceedings of Mensch Und Computer* 2019: 11–20. https://doi.org/10.1145/3340764.3340802.

Sahai, A., E. Anton, E. Stumpf, F. Wefers, and M. Vorlaender. 2012. "Interdisciplinary Auralization of Take-off and Landing Procedures for Subjective Assessment in Virtual Reality Environments." In *18th AIAA/CEAS Aeroacoustics Conference (33rd AIAA Aeroacoustics Conference)*. American Institute of Aeronautics and Astronautics. doi:10.2514/6.2012-2077.

Sanchez, J. 2010. "Recognizing Shapes and Gestures using Sound as Feedback." *CHI '10 Extended Abstracts on Human Factors in Computing Systems* 3063–3306. doi:10.1145/1753846.1753918.

Sawada, Y., Y. Itaguchi, M. Hayashi, K. Aigo, T. Miyagi, M. Miki, T. Kimura, and M. Miyazaki. 2020. "Effects of Synchronised Engine Sound and Vibration Presentation on Visually Induced Motion Sickness." *Scientific Reports* 10 (1): 7553. https://doi.org/10.1038/s41598-020-64302-y.

Schlagowski, R., F. Wildgrube, S. Mertes, C. George, and E. André. 2022. "Flow with the Beat! Human-Centered Design of Virtual Environments for Musical Creativity Support in VR." In *Creativity and Cognition, C&C '22. Association for Computing Machinery*, 428–442. New York, NY. https://doi.org/10.1145/3527927.3532799.

Schoeffler, M., J. L. Gernert, M. Neumayer, S. Westphal, and J. Herre. 2015. "On the Validity of Virtual Reality-Based Auditory Experiments: A Case Study About Ratings of the Overall Listening Experience." *Virtual Reality* 19 (3): 181–200. https://doi.org/10.1007/s10055-015-0270-8.

Scott, C., and M. Medaugh. 2017. "Axial Coding." In *The International Encyclopedia of Communication Research Methods*, 1–2. John Wiley and Sons.

Seabrook, E., R. Kelly, F. Foley, S. Theiler, N. Thomas, G. Wadley, and M. Nedeljkovic. 2020. "Understanding How Virtual Reality Can Support Mindfulness Practice: Mixed Methods Study." *Journal of Medical Internet Research* 22 (3): e16106. https://doi.org/10.2196/16106.

Seele, S., S. Misztal, H. Buhler, R. Herpers, and J. Schild. 2017. "Here's Looking at You Anyway! How Important is Realistic Gaze Behavior in Co-located Social Virtual Reality Games?" In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 531–540. https://doi.org/10.1145/3116595.3116619.

Serafin, S. 2020. "Sonic Interactions in Multimodal Virtual Environments." *Array the Journal of the ICMA*; 16–18. doi:10.25370/array.v20152517.

Serafin, S., M. Geronazzo, C. Erkut, N. C. Nilsson, and R. Nordahl. 2018. "Sonic Interactions in Virtual Reality: State of the Art, Current Challenges, and Future Directions." *IEEE Computer Graphics and Applications* 38: 31–43. https://doi.org/10.1109/MCG.2018.193142628.

Shaw, C., M. Green, J. Liang, and Y. Sun. 1993. "Decoupled Simulation in Virtual Reality with the MR Toolkit." *ACM Transactions on Information Systems* 11 (3): 287–317. doi:10.1145/159161.173948.

Shimamura, R., Q. Feng, Y. Koyama, T. Nakatsuka, S. Fukayama, M. Hamasaki, M. Goto, and S. Morishima. 2020. "Audio–Visual Object Removal in 360-Degree Videos." *The Visual Computer* 36 (10): 2117–2128. https://doi.org/10.1007/s00371-020-01918-1.

Shin, M., S. W. Song, S. J. Kim, and F. Biocca. 2019. "The Effects of 3D Sound in a 360-Degree Live Concert Video on Social Presence, Parasocial Interaction, Enjoyment, and Intent of Financial Supportive Action." *International Journal of Human-Computer Studies* 126: 81–93. https://doi.org/10.1016/j.ijhcs.2019.02.001.

Sikström, E., A. de Götzen, and S. Serafin. 2015. "Wings and Flying in Immersive VR — Controller Type, Sound Effects and Experienced Ownership and Agency." *2015 IEEE Virtual Reality (VR)*, 281–282. https://doi.org/10.1109/VR.2015.7223405.

Sikström, E., E. R. Høeg, L. Mangano, N. C. Nilsson, A. de Götzen, and S. Serafin. 2016a. "Shop 'Til You Hear it Drop: Influence of Interactive Auditory Feedback in a Virtual Reality Supermarket." In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 355–356. https://doi.org/10.1145/2993369.2996343.

Sikström, E., N. C. Nilsson, A. de Götzen, and S. Serafin. 2016b. "Is This Bridge Safe? Evaluation of Audiovisual Cues for a Walk on a Small Bridge Over a Canyon." *2016 IEEE Virtual Reality (VR)*, 285–286. https://doi.org/10.1109/VR.2016.7504765.

Siu, A. F., M. Sinclair, R. Kovacs, E. Ofek, C. Holz, and E. Cutrell. 2020. "Virtual Reality Without Vision: A Haptic and Auditory White Cane to Navigate Complex Virtual Worlds." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376353.

Slater, M. 2003. "A Note on Presence Terminology." *Presence Connect* 3 (3). http://www0.cs.ucl.ac.uk/research/vr/Projects/Presencia/ConsortiumPublications/ucl_cs_papers/presence-terminology.htm.

Slater, M. 2009. "Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1535): 3549–3557. https://doi.org/10.1098/rstb.2009.0138.

Slater, M., C. Gonzalez-Liencres, P. Haggard, C. Vinkers, R. Gregory-Clarke, S. Jelley, Z. Watson, et al. 2020. "The Ethics of Realism in Virtual and Augmented Reality." *Frontiers in Virtual Reality* 1. https://doi.org/10.3389/frvir.2020.00001.

Slater, M., and M. Usoh. 1993. "Representations Systems, Perceptual Position, and Presence in Immersive Virtual Environments." *Presence: Teleoperators and Virtual Environments* 2 (3): 221–233. doi:10.1162/pres.1993.2.3.221.

Sorkin, A., D. Weinshall, and A. Peled. 2008. "The Distortion of Reality Perception in Schizophrenia Patients, as Measured in Virtual Reality." *Studies in Health Technology and Informatics* 132: 475–480.

Souza Silva, W., B. McFadyen, J. Fung, and A. Lamontagne. 2019. "Effects of Age on Obstacle Avoidance While Walking and Deciphering Text Versus Audio Phone Messages." *Gerontology* 65 (5): 524–536. https://doi.org/10.1159/000499067.

Sprung, G., A. Egger, A. Nischelwitzer, R. Strohmaier, and S. Schadenbauer. 2018. "ViReSt—Storytelling with Volumetric Videos." Proceedings of the 11th forum media technology and 4th all around audio symposium. FMT 2018 forum media technology and all around audio symposium, St. Pölten, Austria. http://ceur-ws.org/Vol-2299/paper7.pdf.

Statista. 2022. "Global VR Gaming Market Size 2024." *Statista*. https://www.statista.com/statistics/499714/global-virtual-reality-gaming-sales-revenue/.

Steadman, M. A., C. Kim, J.-H. Lestang, D. F. M. Goodman, and L. Picinali. 2019. "Short-term Effects of Sound Localization Training in Virtual Reality." *Scientific Reports* 9 (1): Article 1. doi:10.1038/s41598-019-54811-w.

Stecker, G. C., T. M. Moore, M. Folkerts, D. Zotkin, and R. Duraiswami. 2018. "Toward Objective Measures of Auditory Co-Immersion in Virtual and Augmented Reality." *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. https://www.aes.org/e-lib/browse.cfm?elib=19668.

Stein, E., and M. M. Goodwin. 2019. "Ambisonics Depth Extensions for Six Degrees of Freedom." Audio Engineering Society Conference: 2019 Aes International Conference on Headphone Technology, August 21. https://www.aes.org/e-lib/browse.cfm?elib=20514.

Steuer, J. 1992. "Defining Virtual Reality: Dimensions Determining Telepresence." *Journal of Communication* 42 (4): 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x.

Suarez, A. S., J. Tissieres, L. S. Vieira, R. Hunter-McHardy, S. K. Sernavski, and S. Serafin. 2017. "A Comparison between Measured and Modelled Head-Related Transfer Functions for an Enhancement of Real-Time 3D Audio Processing for Virtual Reality Environments." In *2017 IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, 1–9. https://doi.org/10.1109/SIVE.2017.7901609.

Svebak, S., and S. Murgatroyd. 1985. "Metamotivational Dominance: A Multimethod Validation of Reversal Theory Constructs." *Journal of Personality and Social Psychology* 48 (1): 107–116. https://doi.org/10.1037/0022-3514.48.1.107.

Tamborini, R., N. D. Bowman, A. Eden, M. Grizzard, and A. Organ. 2010. "Defining Media Enjoyment as the Satisfaction of Intrinsic Needs." *Journal of Communication* 60 (4): 758–777. https://doi.org/10.1111/j.1460-2466.2010.01513.x.

Tennent, P., S. Martindale, S. Benford, D. Darzentas, P. Brundell, and M. Collishaw. 2020. "Thresholds: Embedding Virtual Reality in the Museum." *Journal on Computing and Cultural Heritage* 13 (2): 1–35. https://doi.org/10.1145/3369394.

Togelius, J., G. N. Yannakakis, S. Karakovskiy, and N. Shaker. 2012. "Assessing Believability." In *Believable Bots: Can Computers Play Like People?* edited by P. Hingston, 215–230. Springer. https://doi.org/10.1007/978-3-642-32323-2_9.

Tricco, A. C., E. Lillie, W. Zarin, K. K. O'brien, H. Colquhoun, D. Levac, D. Moher, M. D. J. Peters, T. Horsley, L. Weeks, S. Hempel, E. A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M. G. Wilson, C. Garritty, S. E. Straus. 2018. "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation." *Annals of Internal Medicine* 169 (7): 467–473. doi:10.7326/M18-0850.

Turchet, L., M. Carraro, and M. Tomasetti. 2022. "FreesoundVR: Soundscape Composition in Virtual Reality Using Online Sound Repositories." *Virtual Reality*. [Online]. Accessed October 10, 2022. https://doi.org/10.1007/s10055-022-00705-8.

Turner, S., I. Mcgregor, and P. Turner. 2009. "Listening, Corporeality, Place and Presence." In *Exploration of Space, Technology, and Spatiality: Interdisciplinary Perspectives*, 113–127. IGI Global.

Uebel, K., M. Marselle, A. J. Dean, J. R. Rhodes, and A. Bonn. 2021. "Urban Green Space Soundscapes and Their Perceived Restorativeness." *People and Nature* 3 (3): 756–769. https://doi.org/10.1002/pan3.10215.

Ulsamer, P., K. Pfeffel, and N. H. Müller. 2020. "Brain Activation in Virtual Reality for Attention Guidance." In *Learning and Collaboration Technologies. Human and Technology Ecosystems*, edited by P. Zaphiris and A. Ioannou, 190–200. Springer International Publishing. doi:10.1007/978-3-030-50506-6_14.

Urbanek, M., and F. Güldenpfennig. 2019. "Celebrating 20 Years of Computer-Based Audio Gaming." In *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound, AM'19*, 90–97. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3356590.3356605.

Van den Broeck, W., P. Pamplona, and S. Fernandez Langa. 2017. "Developing a Multi-Device Immersive TV-Experience: User Expectations and Practices." In *Proceedings of the 2nd International Workshop on Multimedia Alternate Realities*, 25–30. https://doi.org/10.1145/3132361.3132362.

Van Elferen, I. 2011. "¡Un Forastero! Issues of Virtuality and Diegesis in Videogame Music." *Music and the Moving Image* 4 (2): 30–39. doi:10.5406/musimoviimag.4.2.0030.

van Elferen, I. 2020. "Ludomusicology and the New Drastic." *Journal of Sound and Music in Games* 1 (1): 103–112. https://doi.org/10.1525/jsmg.2020.1.1.103.

Vosmeer, M., and B. Schouten. 2017. "Project Orpheus a Research Study Into 360° Cinematic VR." In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, 85–90. https://doi.org/10.1145/3077548.3077559.

Watson, D., L. A. Clark, and A. Tellegen. 1988. "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales." *Journal of Personality and Social Psychology* 54 (6): 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063.

Wedoff, R., L. Ball, A. Wang, Y. X. Khoo, L. Lieberman, and K. Rector. 2019. "Virtual Showdown: An Accessible Virtual Reality Game with Scaffolds for Youth with Visual Impairments." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300371.

Westerberg, A., and H. Schoenau-Fog. 2015. "Categorizing Video Game Audio: An Exploration of Auditory-psychological Effects." *Proceedings of the 19th International Academic Mindtrek Conference* 47–54. doi:10.1145/2818187.2818295.

White, K., M. Arntzen, A. Bronkhorst, and M. Meeter. 2014. "Continuous Descent Approach (CDA) Compared to Regular Descent Procedures: Less Annoying?" INTERNOISE 2014 – 43rd international congress on noise control engineering: improving the world through noise control. https://research.vu.nl/en/publications/continuous-descent-approach-cda-compared-to-regular-descent-proce

Wie, O. B., A. H. Pripp, and O. Tvete. 2010. "Unilateral Deafness in Adults: Effects on Communication and Social Interaction." *The Annals of Otology, Rhinology, and Laryngology* 119 (11): 772–781.

Wu, H., D. H. Ashmead, H. Adams, and B. Bodenheimer. 2018. "3D Sound Rendering in a Virtual Environment to Evaluate Pedestrian Street Crossing Decisions at a Roundabout." In *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, 1–6. https://doi.org/10.1109/SIVE.2018.8577195.

Yan, Z., J. Wang, and Z. Li. 2019. "A Multi-Criteria Subjective Evaluation Method for Binaural Audio Rendering Techniques in Virtual Reality Applications." In *2019 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 402–407. https://doi.org/10.1109/ICMEW.2019.00075.

Yin, G., M. J.-D. Otis, P. E. Fortin, J. R. Cooperstock. 2019. "Evaluating Multimodal Feedback for Assembly Tasks in a Virtual Environment." *Proceedings of the ACM on Human-Computer Interaction* 3 (EICS): 21:1-21–11. doi:10.1145/3331163.

Zaini, A. I., R. N. S. R. Yazit, N. A. Goh, Z. A.-A. Gregory, and M. R. Embi. 2022. "Analysis on Recognizability and Legibility of Urban Virtual Environment in Virtual Reality Through Cognitive Maps." In *Advances in Civil Engineering Materials*, edited by M. Awang, L. Ling, and S. S. Emamian, 23–30. Springer. https://doi.org/10.1007/978-981-16-8667-2_4.

Zeroth, J. A., L. M. Dahlquist, and E. C. Foxen-Craft. 2019. "The Effects of Auditory Background Noise and Virtual Reality Technology on Video Game Distraction Analgesia." *Scandinavian Journal of Pain* 19 (1): 207–217. https://doi.org/10.1515/sjpain-2018-0123.

Zhang, B., Z. Yan, J. Wang, Y. Luo, S. Yang, and Z. Fei. 2018. "An Audio-Visual Quality Assessment Methodology in Virtual Reality Environment." In *2018 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 1–6. https://doi.org/10.1109/ICMEW.2018.8551522.

Zhao, J., B. Zhang, Z. Yan, J. Wang, and Z. Fei. 2017. "A Study on the Factors Affecting Audio-Video Subjective Experience in Virtual Reality Environments." In *2017 International Conference on Virtual Reality and Visualization (ICVRV)*, 303–306. https://doi.org/10.1109/ICVRV.2017.00070.