

Review

# Ecological Validity of Immersive Virtual Reality (IVR) Techniques for the Perception of Urban Sound Environments

Chunyang Xu , Tin Oberman , Francesco Aletta , Huan Tong and Jian Kang 

Institute for Environmental Design and Engineering, The Bartlett, University College London, Central House, 14 Upper Woburn Place, London WC1H 0NN, UK; chunyang.xu.18@ucl.ac.uk (C.X.); t.oberman@ucl.ac.uk (T.O.); f.aletta@ucl.ac.uk (F.A.); huan.tong.18@ucl.ac.uk (H.T.)

\* Correspondence: j.kang@ucl.ac.uk

**Abstract:** Immersive Virtual Reality (IVR) is a simulated technology used to deliver multisensory information to people under different environmental conditions. When IVR is generally applied in urban planning and soundscape research, it reveals attractive possibilities for the assessment of urban sound environments with higher immersion for human participation. In virtual sound environments, various topics and measures are designed to collect subjective responses from participants under simulated laboratory conditions. Soundscape or noise assessment studies during virtual experiences adopt an evaluation approach similar to in situ methods. This paper aims to review the approaches that are utilized to assess the ecological validity of IVR for the perception of urban sound environments and the necessary technologies during audio–visual reproduction to establish a dynamic IVR experience that ensures ecological validity. The review shows that, through the use of laboratory tests including subjective response surveys, cognitive performance tests and physiological responses, the ecological validity of IVR can be assessed for the perception of urban sound environments. The reproduction system with head-tracking functions synchronizing spatial audio and visual stimuli (e.g., head-mounted displays (HMDs) with first-order Ambisonics (FOA)-tracked binaural playback) represents the prevailing trend to achieve high ecological validity. These studies potentially contribute to the outcomes of a normalized evaluation framework for subjective soundscape and noise assessments in virtual environments.

**Keywords:** ecological validity; virtual reality; sound environment; soundscape; sound environment perception



**Citation:** Xu, C.; Oberman, T.; Aletta, F.; Tong, H.; Kang, J. Ecological Validity of Immersive Virtual Reality (IVR) Techniques for the Perception of Urban Sound Environments.

*Acoustics* **2021**, *3*, 11–24.

<https://dx.doi.org/10.3390/acoustics3010003>

Received: 17 November 2020

Accepted: 23 December 2020

Published: 25 December 2020

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Ecological validity was introduced in the 1980s to evaluate the outcomes of a laboratory experiment focused on visual perception [1]. Ecological validity describes the degree to which results obtained in a controlled laboratory experiment are related to those obtained in the real world [2]. The discussion of the ecological approach regarding its internal validity and experimental control began in the 1980s with cognitive and behavioral psychology research [1,3], and these two factors are still significant factors in the design and undertaking of an ecological approach study. Under laboratory conditions, researchers should give participants corresponding environmental cues and instructions to enable the reactivation of the cognitive processes of participants that were determined in actual situations [4]. For high ecological validity, the findings in the laboratory can be generalized into real-life settings [2]. As a simulated technology, Immersive Virtual Reality (IVR) places the user inside an experience, which allows the impact on participants of a new environment with complex social interactions and contexts to be assessed [5–7]. In 2001, Bishop et al. [8] reported their non-IVR assessments of path choices on a country walk, and they agreed that faster computers and better display systems make the virtual environment experience more credible. Thus, low ecological validity resulting from non-sufficient immersiveness could be a limiting factor for the generalizability of data collected from laboratory experiments.

The need for more research that explores applications of perceptual simulations in general and related questions of validity and reliability has been stressed ever since the emergence of environmental simulation as a research paradigm.

Ecological validity has been conceptualized into two approaches: verisimilitude and veridicality. Verisimilitude refers to the extent of similarity of a virtual experience to relevant environmental behaviors [9]; it reflects the similarity of the task demands between the test in the laboratory and the real world [10]. This approach attempts to create new evaluation assessments with ecological goals [11]. Veridicality refers to the degree of accuracy in predicting some environmental behaviors [12,13]; the establishment of veridicality is required to assess the results from the laboratory test and the measures in the real-world. There are some limitations for both approaches. One limitation of the veridicality approach is that, for those conditions which are not likely to be reproduced in the real world or that have a high cost, the outcomes from real-world measures cannot correlate with experiment results. When using the verisimilitude approach alone, no empirical data are needed to claim that the evaluation is similar to real life settings [11].

Virtual reality has revealed a functional rapprochement that fuses the boundary between the laboratory and real life [5]. Through multisensory stimuli with experimental control, participants tend to respond realistically to virtual situations as if they were in a real environment [14–17]. The responses to a virtual environment are generated when place illusion (PI) and plausibility illusion (Psi) occur at the same time [14–16]. The ecological approach studies based on virtual reality provide controlled dynamic presentations of background narratives to enhance the affective experience and social interactions [3,18]. From a methodological viewpoint, environmental conditions and test results can be ecologically validated through virtual reality technologies according to a subjective evaluation framework. Numerous researchers have examined ecological validity in different topics and fields with the comparison of a virtual environment and real life [19–23].

Spatial audio is a technique of creating sound in a 3D space; then, a listener can hear the sound from any direction in a sphere [24]. Because of this feature, it is often combined with virtual reality to render auditory stimuli. For the auralization of spatial audio, head-tracking requires the reproduction of a dynamic sound field based on the real-time position of the head within Euclidean space. Binaural recordings only reproduce the sound field of both ears at the time of recording, which shows the incompatibility between binaural recordings and head-tracking during auralization. Ambisonics is a sound reproduction technique used for recording and playing-back spatial audio, and it is based on the spherical harmonic decomposition of the sound field [25]. Ambisonics enables a listener to experience a spatially-accurate perception of the sound field [26], and this reproduction technique was originally introduced by Gerzon [27]. In the case of first-order Ambisonics (FOA)—currently the most widespread Ambisonics recording technique—the signals are recorded as four audio channels, most usually in the so-called A-format. The audio files needed to reproduce such recordings are known as B-format audio files, which are converted from the A-format. The B-format can be decoded into any speaker array matching the needs of dynamic auralization under Immersive Virtual Reality (IVR), including higher-order Ambisonics (HOA). HOA has a higher spatial resolution based on higher-order spherical harmonics [25]. Head-related transfer function (HRTF) is considered as a frequency response describing the sound pressure transformation from a free field point source to the eardrum [28]. When this filtering is not applied with the listener's own HRTF (acknowledging individual head size, auricle size and shape, etc.) [29–31], front-back and elevation confusions in localization typically occur [32].

For many urban sound assessment studies, in situ surveys have been widely applied as a conventional method to evaluate certain sound environments [5,33,34]. In soundscape or noise assessment studies, researchers expect the presentation of controlled experimental conditions to participants; e.g., recorded audio and reconstructed visual stimuli in a listening room. Therefore, researchers introduced laboratory tests to validate their research

questions with human participation. All simulations under laboratory conditions attempt to represent some aspects of the environment as accurately as possible to assess human responses. In urban noise prediction and soundscape assessment research, an audio–visual system is a conventional and valid approach to render essential information or cues during human participation. The audio–visual interaction influences the perception of the soundscape and global environment, as shown in previous studies [35–41]. For interior spaces with VR techniques, several studies have assessed the evaluation of indoor noise protection with head-mounted displays (HMDs) [42], the main uses of auralization [43], the influence of visual distance [44], the use of water features [45] and the spatial representations of visually impaired participants [46,47]. The urban sound environment in this review refers primarily to sound sources originating outdoors or in urban public spaces, and it reflects, to some extent, the mobility of people and the multifunctionality of urban spaces.

The evaluated multisensory method shows enormous significance in helping participants to perceive environments holistically. The reproduction system of listening tests needed to be adapted to the purpose of the study to allow the subjects to treat the test samples as potentially familiar experiences through cognitive processes elaborated in actual situations. With the aid of immersive virtual reality, the installation of laboratory conditions was performed with the aim of reproducing urban sound environments and presenting a multisensory experience to participants. A subjective test of immersive virtual reality reproduction in urban sound environment assessments would show high veridicality if it correlated well with measures of perceptual responses in the real world.

The concept of ecological validity has been extended from psychological experiments to the domain of complex sound environment perception. It is not only related to the evaluation methods during laboratory tests, but also closely associated with the developing IVR technologies. Attempting to establish a standardized soundscape evaluation protocol with high veridicality under an immersive virtual environment has a broader impact on the practice of soundscape planning and design. The research on soundscape standardization has discussed the definitions, variety of contexts, evaluation methods and reporting requirements [48,49].

The International Organization for Standardization Technical Standard (ISO TS) 12913-2 [34] introduced two common recording techniques in soundscape research: binaural and Ambisonics. The standard states that if some environmental factors are not present or differ during playback, the outcomes could possibly result in different impressions to those received in the original context. In terms of the statement of ISO TS 12913-2 [34], the validity of these auralization techniques combined with other environmental factors still presents some uncertainty. The ISO TS 12913-3 [50] stated that the key factors to consider when conducting ecologically valid laboratory studies are the effect of memory, the duration of exposure to each of the stimuli and the auditory immersiveness. As a multisensory tool, IVR could deliver more environmental stimuli than conventional 2D rendering methods. A comparison of the ecological validity using IVR for urban sound environments with different reproduction techniques and research topics was made. In this review, we aim to investigate (1) which kinds of approaches can be utilized and integrated to assess the ecological validity of IVR when humans perceive urban sound environments and (2) which technologies are necessary during audio–visual reproduction to establish a dynamic IVR system to assess the perception of urban sound environments.

## 2. Methodologies

### 2.1. Search Strategy and Eligibility Criteria

There were no pre-defined protocol registrations for this review. The basic process and data extraction forms were agreed upon at the beginning of the review work. The study was performed under the guidance of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [51].

Given the exploratory nature of this study, as many studies do not directly mention “ecological validity”, and they may not include the terms “ecologically valid”, “ecologically

validate” and similar expressions, the studies were selected manually according to the following inclusion criteria: (1) original participatory studies using virtual reality techniques conducted in a laboratory, and (2) studies collecting subjective responses under virtual environments. The subject areas included “ecological validity”, “acoustics” and “virtual reality”. Some studies did not directly mention “ecological validity”, but the workflow was under the framework of the virtual sound environment evaluation described above. These studies were selected in the review with full-text scanning.

Studies were identified by searching the electronic database, scanning reference lists of articles and in consultation with experts in the field. At least one search database should be used in systematic reviews [51], and a literature search was conducted using the Web of Science. Only peer-reviewed journal articles published in English were considered. This search was applied to the Web of Science (1980–2020). The last search was run on 01 July 2020. We used the following search terms to search the databases: “sound”, “perception”, “participant” and “virtual reality”. An eligibility assessment of the review was performed independently in an un-blinded standardized manner by one researcher.

## 2.2. Data Extraction

Information was extracted from each included document regarding (1) the research focus of the studies, (2) participant numbers, (3) in situ responses vs. laboratory experiment data and (4) the main parameters selected in the studies. Considering ecological validity across the selected studies with various topics and different outcomes, a qualitative approach was adopted to answer the review questions.

## 3. Results

The initial results showed 65 documents. Fifty-three items were excluded because the topic of the papers failed to meet the eligibility criteria, which included (1) the studies not using a virtual reality head-mounted device, and (2) the studies not involving sound-related perception. The full texts of the remaining 12 papers were accessed, and these 12 papers were included in the review.

### 3.1. Ecological Validity with Subjective Responses

Table 1 shows the research focus, the participant numbers, in situ responses vs. experiment data, the main parameters and variables in these studies. These studies with IVR had different emphases on their subjective evaluation and research focuses, and they assessed ecological validity with subjective responses varying from environmental preferences/quality, audio/visual indicators, coupled interactions and reproduction quality.

Generally, these studies were not only limited to one topic, and several topics were often integrated together. The audio–visual interaction was also one of the sub-topics of these studies. Most of these works addressed the importance of audio–visual interaction in IVR-based soundscape or noise assessments. The audio–visual interaction in these studies was discussed in an attempt to interpret how participants perceived the virtual environments, and the ecological validity was also tested with their research questions. Global environmental evaluation, visual and acoustic coherence and familiarity and visual and acoustic congruence were compared, respectively, through the field survey and the laboratory experiment to jointly validate the acoustic and visual congruence between the simulated and real world [5]. Both groups in the in situ session and the laboratory session with 16 participants, respectively, were recruited for the in situ and laboratory sessions. Both comparison groups showed robust similarity in visuo-acoustic coherence and familiarity and the visuo-acoustic salience of urban, human and natural activities.

Related to audio–visual interactions, in 2020, Jeon and Jo [52] examined the contribution of audio and visual stimuli in the evaluation of urban environment satisfaction under an immersive virtual environment. Three conditions were considered: (1) audio only, (2) vision only and (3) audio–visual interaction. The contributions of audio and visual information on overall satisfaction were 24% and 76%, respectively. The study by Ruotolo

---

et al. in 2013 [15] asked the participants to answer questions about auditory and visual annoyance, respectively. The results presented in their study showed both auditory information and visual information in a close interaction, supporting participants perceiving the virtual environment holistically. Aletta et al. in 2016 [53] carried out a study to investigate the chiller noise involved with the distance to a source and the visibility of a source, as well as introduction—or not—of a visual reference context and performance—or not—of a cognitive task. They found that the visibility of a source is not a significant influencing factor for noise perception for the kinds of chillers examined in the study.

**Table 1.** Virtual environment evaluation studies with subjective responses.

References	Research Focus	Participant Number	In Situ Responses vs. Experiment Data	Auditory Perception	Visual Perception	Coupled or Other Variables
(Iachini et al., 2012) [54]	Acoustic comfort aboard metros	51	x	Annoyance	–	–
(Ruotolo et al., 2013) [15]	Noise assessment for a motorway	20	x	Audio annoyance	Visual annoyance	–
(Maffei, Iachini, et al., 2013) [55]	The effects of visual and acoustical aspects of the impact of a wind farm in a quiet area	46	x	Perceived attributes	Visual Pleasantness	General environment evaluation
(Maffei, Masullo, et al., 2013) [56]	The influence of visual characteristics of barriers on railway noise perception	41	x	Annoyance, perceived loudness	Visual pleasantness	–
(Aletta et al., 2016) [53]	The effect of vision on the perception of the chiller noise	26	x	Perceived loudness, noise annoyance	Visual unpleasantness	–
(Maffei et al., 2016) [5]	Global sound environmental quality	16 in situ, 16 in the laboratory	✓	Acoustic coherence, and familiarity	Visual coherence, and familiarity	Global qualitative evaluations
(Sanchez et al., 2017) [57]	The role of noise in the audio-visual design of an urban public space	71	x	–	–	The audio-visual interaction for preference and reality evaluation
(Yu et al., 2017) [58]	Noise and visual intrusion from wind parks on affective and cognitive performances	20	x	Audio annoyance	Visual annoyance	–
(Hong et al., 2019) [59]	The FOA reproduction comparison with in situ soundscape evaluation	5, 12, 13 in three days (in situ), the same participants in the laboratory	✓	Overall soundscape quality	Perceived spatial quality	Distinctiveness, immersiveness, realism, reproduction fidelity
(Jeon and Jo, 2019) [60]	Road traffic noise perception in urban high-rise residential buildings	40	x	Perceived loudness, annoyance, sound acceptance	–	Perceived distance, perceived directionality, perceived width, immersion, realism, and perceived externalization
(Sun et al., 2019) [61]	Classifying soundscapes of urban public open spaces	20 for Group 1, 20 for Group 2	x	Classification components, psycho-acoustical indicators, and saliency	Visual factors	–
(Jeon and Jo, 2020) [52]	The relationship between overall satisfaction of the urban environment and audio-visual interactions	30	x	Sound preference, soundscape attributes	Visual preference, visual attributes	Environment satisfaction



In 2019, Jeon and Jo [60] carried out a study to assess the noise in urban high-rise residential buildings. They reported that the directional and visual information generated by HRTFs and HMDs could affect sound perception and virtual environmental immersion. Two parameters, HRTFs and HMDs, were coupled into four cases: no HRTF–no HMD, no HRTF–HMD, HRTF–no HMD and HRTF–HMD. The results showed that the contribution of the HRTF to subjective responses was 77% higher than the contribution of HMD, at 23%. This study showed the applicability and necessity of the HRTF and HMD to assess noise in terms of the audio–visual interactions under immersive virtual environments. In 2012, Iachini et al. [54] assessed acoustic comfort aboard metros through subjective annoyance and cognitive performance measures. In their findings, visual contexts could be considered as a modulating method affecting noise annoyance for people. Noise barrier designs are generally associated with noise assessment, and different noise barrier designs were assessed under an immersive virtual environment [56,57]; different project solutions concerning noise mitigation in order to obtain more reliable results on local residents were also examined.

The potential environmental risks and negative effects of wind parks, as emerging landscape projects, were also evaluated under virtual environments. Maffei, Iachini, et al. in 2013 [55] stated that the noise perception of wind turbines under immersive virtual reality requires extended experiments to ensure its ecological validity, especially the results from in situ sessions. In 2017, Yu et al. [58] conducted a subjective test revealing that wind parks can increase both the aural and visual annoyance associated with personal attitudes toward wind parks. The research of virtual reality technologies in the sound environments of wind parks ecologically validated these potential negative influences.

Soundscape evaluations show a trend of using multi-dimensional attributes to test participants' perception in a virtual environment. In 2019, Hong et al. [59] carried out a study exploring the ecological validity of reproduced acoustic environments based on three spatial audio reproduction methods. The main indicators in their study included sound preferences, visual preferences, soundscape attributes, visual attributes and environmental satisfaction, as shown in Table 1. In 2019, Sun et al. [61] proposed a hierarchical soundscape classification method using virtual reality playback with a participatory experiment inside a soundproof booth. The method, based on different classification components, could be potentially validated by verification on an independent dataset.

In immersive virtual reality laboratory experiments, the numbers of participants differ in different subjective studies. The minimum number of participants in the subjective test was 16, in the work by Maffei et al. in 2016 [5], and the maximum number reached 71, in the work by Sanchez et al. in 2017 [57]. The number of participants for most subjective tests in the laboratory ranges from 20 to 60.

### 3.2. Reproduction Systems

The reproduction systems in these studies mainly include two aspects of auralization and visualization, as shown in Table 2. To simulate an immersive auditory environment, Ambisonics is a prevailing method to record and auralize sounds, which allows various decoding patterns with the flexibility to lay out loudspeaker positions or headphones. In the headphone-based reproduction method, the recorded stimuli captured in Ambisonics formats are most usually presented as either head-tracked or static binaural renderings [52,59]. In the loudspeaker array-based reproduction method, there is no need for software to compensate for head movement in real time [59]. The spatial recognition is influenced by the use of the HMD, and the impact of perceived realism will significantly increase compared with the condition without HMD [59]. Simulated visual environments can also be built using software including—but not limited to—3ds Max [56,57], Google SketchUp [5,15,54,60], Unity [57,58], Kubity [60] and WorldViz [53–55]. In 2019, Hong et al. [59] reported no significant differences in perceived dominant sound sources and affective soundscape quality between reproduction and in situ results. These findings are in agreement with

previous studies showing that IVR HMDs with Ambisonics could be a reliable tool for soundscape assessment as an alternative to in situ surveys.

Some devices have been introduced to record information and render stimuli. A panoramic camera is usually used to record omnidirectional videos as visual stimuli in the laboratory test [52,59]. A hybrid and simultaneous audio and video recording setup was used in the study by Sun et al. in 2019 [61]. This setup consists of binaural audio (an artificial head with windshield and binaural recording device), a FOA microphone and a 360° video camera. A mobile device (a Google Cardboard headset) was also used in the evaluation of the audio–visual perception of wind parks. This portable HMD also showed the potential to provide an immersive experience in response to participants' head movements.

**Table 2.** Auralization and visualization during the participatory experiments. HMD: head-mounted display.

<b>Auralization</b>	
Recordings	Playback
Binaural audio signal recordings [15,53–56,58]	Headphones [15,52,54,57,58,60,61]
Ambisonics recordings [5,52,57,59,61]	A number of loudspeakers, and a sub-woofer [55,56]
	Headphones with a sub-woofer [53]
	5.1-format loudspeaker configuration [5]
<b>Visualization</b>	
Visual construction methods	Visual rendering
3ds Max [56,57]	HMD [5,15,42,52–61]
Google SketchUp [5,15,54,60]	
WorldViz [53–55]	
Unity [57,58]	
Kubity [60]	
Panoramic views [52,58,60,61]	

Notably, owing to the fact that the entire IVR industry is driven by both hardware and software upgrades, older ecological validity studies on virtual environments face limitations in terms of their utility or efficacy. It would be expected that the advancement in the computation of IVR simulations would ultimately increase the ecological validity of participatory studies conducted in laboratories. A comparison of the technical parameters of all IVR systems in these studies shows the limitations of initial research and how these limitations are gradually improved by subsequent studies. However, due to the lack of control measures across the analyzed studies, it was not possible to conduct such a comparison. We cannot systematically assess the differences between the studies.

## 4. Discussion

### 4.1. Subjective Response, Cognitive Performance and Physiological Response

Many studies have suggested that urban noise can negatively affect people's cognitive functions and influence their daily life [62–64]. Subjective responses may not show annoyance regarding urban noise, but the cognitive performance may be affected. Thus, during the laboratory test, some studies also used cognitive tasks to evaluate the cognitive performance caused by the virtual environment [15,53,54,58]. Related to stress recovery, researchers have used measures based on the physiological responses of participants. Annerstedt et al. in 2013 [65] conducted a study to investigate the sounds of nature inducing physiological stress recovery, and the Trier Social Stress Test (TSST), as a highly standardized protocol for inducing stress, was applied in their study. Cortisol, heart rate, T-wave amplitude (TWA), and heart rate variability (HRV) were tested to analyze the physiological stress recovery induced by the sounds of nature. Hedblom et al. in 2019 [66] adopted mild electrical shocks and skin conductance measurements to evaluate the stress recovery under virtual environments with a birdsong–traffic noise interaction. Compared with subjective responses, physiological responses do not directly reflect the



relationship between subjective sound preferences and characteristics of acoustic environments. Thus, these three methods can jointly assess the ecological validity of complex sound environment perception.

#### 4.2. Other Visual Rendering Methods

For visual rendering, many studies used non-HMD options. Some of them adopted non-immersive methods, such as a monitor screen [8,19,45,67–69], visual screen [70] and 2D projection [71,72]. Some of the studies utilized the immersive Cave Automatic Virtual Environment (CAVE) system [65,73]. The CAVE system was first introduced in 1992 [74], and the aim of its invention was to provide a one-to-many visualization experience that utilizes large projection screens [75]. Compared with a CAVE system, HMD has some problems, especially when one user is trying to interact with other users, and it does not offer interaction with real objects aside from VR control devices [76]. The large footprint, the cost of high-resolution projectors and the human–computer interaction are also reported to be limitations for a CAVE system [76].

Studies without visual stimuli were also conducted [4,77,78]. A visual component presents rationality when examining the ecological validity of auditory perception. The coupled audio–visual interaction is associated with the spatial attributes of sound perception—e.g., distance, width and directionality [60]—and it also provides an animated visual anchor, improving the sense of presence and immersiveness during the subjective evaluation [77,79].

#### 4.3. Verisimilitude and Veridicality

Verisimilitude and veridicality in IVR-based sound environment research have different emphases according to their definitions. Establishing verisimilitude and veridicality in a subjective evaluation experiment allows a virtual sound environment to be perceived with reliable ecological validity. The IVR research involved with verisimilitude in soundscape or noise assessments assumes that the stimuli in the test and the cognitive processing are sufficiently similar to the psychological construct of corresponding scenarios in the real world. The verisimilitude approach is likely to focus on specific tasks in the laboratory test similar to the task demands in the real world. The evaluation indicators and questionnaire design can be formatted in a quite similar way to a participatory experiment. Sanchez et al. in 2017 [57] pointed out that their study did not strictly prove that audio–visual designs in a virtual environment would lead to the predicted pleasantness of real environments. Establishing verisimilitude in soundscape evaluation is more intuitive compared with establishing a new cognitive task or a clinic neuropsychological assessment. However, when researchers discuss the relationship between subjective responses, cognitive performance and physiological responses, they need to carefully examine the verisimilitude approach with which some aspects of testing conditions limit the applicability of a method without empirical data to the real world.

A few studies validated veridicality in IVR-based soundscape or noise assessments. The pioneering studies examined several fundamental playback systems. In 2005, Guastavino et al. [4] explored the linguistic analysis of verbal data in soundscape reproduction through a field survey and two listening tests. Both listening tests compared exposure to the stimuli reproduced via stereophonic and Ambisonics approaches. They pointed out that both neutral visual elements and a good sense of spatial immersion should be provided to ensure ecological validity when testing the effects of urban background noise. Both reproduction methods have been demonstrated to be ecologically valid tools in terms of source identification. However, IVR was not applied in their study. Many perceptual attributes and indicators have been selected to describe the similarity between the real world and the laboratory conditions. In 2016, Maffei et al. [5] compared the congruence between audio and visual elements, and there was no significant difference in the perceived global quality of the environments in both the simulated and real world in their results. The global quality of the environments was shown to have high veridicality under the framework of subjective evaluation. The findings are consistent with the results of audio–visual

interaction evaluation studies conducted in urban sound environments. In 2019, Hong et al. [59] validated three Ambisonics reproduction methods and tested their veridicality under a virtual sound environment related to the performance of reproduction methods. Immersive virtual reality has been shown to be a valid tool to simulate multisensory environments not only by acousticians but also in clinical neuroscience, cognitive psychology and other research fields. When researchers adopt the verisimilitude approach, they believe that the reproduction system and the subjective test have veridicality. In addition, there are also some difficulties to validate veridicality resulting from the complex contexts and unpredictability of outdoor sound environments. For outdoor sound environments, it is sometimes impossible to measure the real-world; e.g., a projected area without construction. Some contextual conditions cannot be changed independently in the real world as well.

It is notable that two studies addressed realism in their subjective experiments. The study by Jeon and Jo in 2019 [60] validated that the usage of HMD significantly increased the impact on the recognition of realism. In 2019, Hong et al. [59] conducted both in situ and laboratory experiments to assess the performance of different Ambisonics reproduction systems in perception. They both successfully assessed realism in their studies. The former de-emphasized the verisimilitude to the real world, and they underlined the realism difference brought by HMD compared with the non-HMD condition. The latter conducted a veridicality study with in situ responses, and they described the degree to which different reproduction approaches were similar to reality. When both verisimilitude and veridicality are examined, the most ecologically valid studies [5,59] revealed the congruence between immersive virtual experience and real experience along with multisensory stimuli.

#### 4.4. Limitations

An IVR system in soundscape or noise assessment should be adapted to the relationship between human cognition and subjective perception during the laboratory experiment. The diversity of IVR rendering techniques also brings an unnormalized experience to participants. An online survey has been introduced as a non-IVR tool to evaluate soundscape and noise perception [68]. Web-based virtual reality was constructed in computationally cheap ways, and it could be improved with higher auralization and visualization quality. The one-to-one nature of tests also showed that the laboratory test cannot reach the sample size of traditional surveys. More economical and vivid reproducing systems following the development in hardware and software show higher veridicality.

HRTFs significantly contribute to localization performance [80,81]; e.g., sound recognition of the direction and width of the source [60]. Compared with the non-HRTF environment, the results of the subjective responses of immersion, realism and externalization are higher in the HRTF case [60]. Individualized and non-individualized HRTFs were utilized to assess various perceptual attributes by Simon et al. in 2016 [31]. It is necessary to select a suitable HRTF that is well matched to the listener's own HRTF [31] to ensure ecological validity in terms of sound source localization, and it can be an individualized HRTF or from a HRTF database. For different sound environments, such as a lively urban square with multiple water features, a quiet park or a park adjacent to a motorway, whether sound source localization is considered a key feature or not [82,83], the choice of an HRTF could differ in terms of ecological validity, and further studies are still needed.

At the moment, a head-tracking display system synchronizing FOA-tracked binaural playback shows reliable validity under immersive virtual experiences for complex sound environment perception. Compared with FOA, HOA significantly improves the quality of this experience [84]. Different systems of HOA have already been implemented as hearing aids research for subjects with hearing loss [85,86]. HOA is becoming popular in industrial applications such as Youtube360 and Facebook360 [87], and it shows great potential for the ecological validity of IVR in further urban sound environment studies.

## 5. Conclusions

This paper aims to review the approaches to assess the ecological validity of IVR for the perception of urban sound environments and the necessary technologies during audio-visual reproduction ensuring ecological validity. The review qualitatively shows that immersive virtual reality techniques have the potential to contribute greatly as an ecologically valid tool in soundscape or noise assessments. The ecological validity of virtual reality to assess urban sound environments is multimodal, dynamic and contextual. The main conclusions of this work are as follows.

1. Through the approaches of laboratory tests including subjective response surveys, cognitive performance tests and physiological responses, the ecological validity of complex sound environment perception can be assessed for IVR. With participatory experiments in situ and in a laboratory, the veridicality of IVR can be verified through subjective responses including environmental preferences/quality, audio-visual indicators (e.g., pleasantness and annoyance), coupled interactions and reproduction quality (e.g., realism and immersiveness).
2. A head-tracking unit with a display and synchronized spatial audio (e.g., HMD with FOA-tracked binaural playback) is advantageous to assess ecological validity in immersive virtual environments. When the urban sound environment research involves interaction among multiple users, a CAVE system should be considered. With higher spatial resolutions, HOA also shows increasing potential for the ecological validity of IVR in urban sound environment research.

These studies on ecological validity with the utilized evaluation methods also go beyond the outcomes gained towards a normalized framework in soundscape and noise assessment protocols. For standardized soundscape evaluation, the ISO TS 12913 series should give more detailed guidelines and specifications on the establishment of an IVR system. In particular, to deliver a dynamic virtual experience, more research is needed on the influence of the Ambisonics orders of complexity at the recording and reproduction stages, and issues such as encoding and decoding Ambisonics formats, on soundscape perception. The pursuit of a standardized soundscape evaluation protocol and IVR-based soundscape research can serve to enhance the field as a whole.

**Author Contributions:** All authors of this research contributed to its conceptualization and writing—review and editing. Methodology, C.X., H.T.; formal analysis, C.X., T.O., F.A.; writing—original draft preparation, C.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded through the European Research Council (ERC) Advanced Grant (no. 740696) on “Soundscape Indices” (SSID).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No datasets were generated or analyzed during the current study.

**Acknowledgments:** The authors would like to gratefully acknowledge the support of the following people: Mengting Liu (UCL), and Ryan Bellman (UCL).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gibson, J.J. *The Ecological Approach to Visual Perception: Classic Edition*; Psychology Press: Hove, UK, 2014. [[CrossRef](#)]
2. Naugle, R.I.; Chelune, G.J. Integrating neuropsychological and “real-life” data: A neuropsychological model for assessing everyday functioning. In *The Neuropsychology of Everyday Life: Assessment and Basic Competencies*; Springer: Berlin, Germany, 1990; pp. 57–73.
3. Parsons, T.D. Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front. Hum. Neurosci.* **2015**, *9*, 660. [[CrossRef](#)]
4. Guastavino, C.; Katz, B.F.; Polack, J.-D.; Levitin, D.J.; Dubois, D. Ecological validity of soundscape reproduction. *Acta Acust. United Acust.* **2005**, *91*, 333–341.

5. Maffei, L.; Masullo, M.; Pascale, A.; Ruggiero, G.; Romero, V.P. Immersive virtual reality in community planning: Acoustic and visual congruence of simulated vs real world. *Sustain. Cities Soc.* **2016**, *27*, 338–345. [[CrossRef](#)]
6. Kotz, S.; Iachini, T.; Coello, Y.; Frassinetti, F.; Ruggiero, G. Body Space in Social Interactions: A Comparison of Reaching and Comfort Distance in Immersive Virtual Reality. *PLoS ONE* **2014**, *9*, e111511. [[CrossRef](#)]
7. Loomis, J.M.; Blascovich, J.J.; Beall, A.C. Immersive virtual environment technology as a basic research tool in psychology. *Behav. Res. Methods Instrum. Comput.* **1999**, *31*, 557–564. [[CrossRef](#)]
8. Bishop, I.D.; Wherrett, J.R.; Miller, D.R. Assessment of path choices on a country walk using a virtual environment. *Landsc. Urban Plan.* **2001**, *52*, 225–237. [[CrossRef](#)]
9. Spooner, D.M.; Pachana, N.A. Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically intact populations. *Arch. Clin. Neuropsychol.* **2006**, *21*, 327–337. [[CrossRef](#)]
10. Sbordone, R.J.; Long, C. *Ecological Validity of Neuropsychological Testing*; CRC Press: Boca Raton, FL, USA, 1996.
11. Chaytor, N.; Schmitter-Edgecombe, M. The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychol. Rev.* **2003**, *13*, 181–197. [[CrossRef](#)]
12. Kenworthy, L.; Yerys, B.E.; Anthony, L.G.; Wallace, G.L. Understanding executive control in autism spectrum disorders in the lab and in the real world. *Neuropsychol. Rev.* **2008**, *18*, 320–338. [[CrossRef](#)]
13. Wood, R.L.; Lioffi, C. The ecological validity of executive tests in a severely brain injured sample. *Arch. Clin. Neuropsychol.* **2006**, *21*, 429–437. [[CrossRef](#)]
14. Jallouli, J.; Moreau, G. An immersive path-based study of wind turbines' landscape: A French case in Plouguin. *Renew. Energy* **2009**, *34*, 597–607. [[CrossRef](#)]
15. Ruotolo, F.; Maffei, L.; Di Gabriele, M.; Iachini, T.; Masullo, M.; Ruggiero, G.; Senese, V.P. Immersive virtual reality and environmental noise assessment: An innovative audio-visual approach. *Environ. Impact Assess. Rev.* **2013**, *41*, 10–20. [[CrossRef](#)]
16. Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* **2009**, *364*, 3549–3557. [[CrossRef](#)]
17. Oberman, T.; Šćitaroci, B.B.O.; Jambrošić, K. Towards a Virtual Soundwalk. In *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*; IGI Global: Hershey, PA, USA, 2018; pp. 317–343.
18. Diemer, J.; Alpers, G.W.; Peperkorn, H.M.; Shibani, Y.; Mühlberger, A. The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Front. Psychol.* **2015**, *6*, 26. [[CrossRef](#)]
19. Bishop, I.D.; Rohrman, B. Subjective responses to simulated and real environments: A comparison. *Landsc. Urban Plan.* **2003**, *65*, 261–277. [[CrossRef](#)]
20. Chamilothori, K.; Wienold, J.; Andersen, M. Adequacy of immersive virtual reality for the perception of daylight spaces: Comparison of real and virtual environments. *Leukos* **2019**, *15*, 203–226. [[CrossRef](#)]
21. Higuera-Trujillo, J.L.; Maldonado, J.L.-T.; Millán, C.L. Psychological and physiological human responses to simulated and real environments: A comparison between Photographs, 360 Panoramas, and Virtual Reality. *Appl. Ergon.* **2017**, *65*, 398–409. [[CrossRef](#)]
22. Kort, Y.A.D.; Ijsselstein, W.A.; Kooijman, J.; Schuurmans, Y. Virtual laboratories: Comparability of real and virtual environments for environmental psychology. *Presence Teleoperators Virtual Environ.* **2003**, *12*, 360–373. [[CrossRef](#)]
23. Rossetti, T.; Hurtubia, R. An assessment of the ecological validity of immersive videos in stated preference surveys. *J. Choice Model.* **2020**, *34*, 100198. [[CrossRef](#)]
24. Rumsey, F. *Spatial Audio*; Taylor & Francis: Oxfordshire, UK, 2001.
25. Daniel, J.; Moreau, S.; Nicol, R. *Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging*; Audio Engineering Society Convention 114; Audio Engineering Society: New York, NY, USA, 2003.
26. Tylka, J.G.; Choueiri, E.Y. Models for Evaluating Navigational Techniques for Higher-Order Ambisonics. In *Proceedings of the Meetings on Acoustics 173EAA*; Acoustical Society of America: Washington, DC, USA, 2017; p. 050009.
27. Gerzon, M.A. Ambisonics in multichannel broadcasting and video. *J. Audio Eng. Soc.* **1985**, *33*, 859–871.
28. Gardner, W.G.; Martin, K.D. HRTF measurements of a KEMAR. *J. Acoust. Soc. Am.* **1995**, *97*, 3907–3908. [[CrossRef](#)]
29. Seeber, B.U.; Fastl, H. *Subjective Selection of Non-Individual Head-Related Transfer Functions*; Georgia Institute of Technology: Atlanta, GA, USA, 2003.
30. Wenzel, E.M.; Arruda, M.; Kistler, D.J.; Wightman, F.L. Localization using nonindividualized head-related transfer functions. *J. Acoust. Soc. Am.* **1993**, *94*, 111–123. [[CrossRef](#)]
31. Simon, L.S.; Zacharov, N.; Katz, B.F. Perceptual attributes for the comparison of head-related transfer functions. *J. Acoust. Soc. Am.* **2016**, *140*, 3623–3632. [[CrossRef](#)]
32. Tan, C.-J.; Gan, W.-S. User-defined spectral manipulation of HRTF for improved localisation in 3D sound systems. *Electron. Lett.* **1998**, *34*, 2387–2389. [[CrossRef](#)]
33. Cadena, L.F.H.; Soares, A.C.L.; Pavón, I.; Coelho, L.B. Assessing soundscape: Comparison between in situ and laboratory methodologies. *Noise Mapp.* **2017**, *4*, 57–66. [[CrossRef](#)]
34. ISO. *TS 12913-2: 2018—Acoustics—Soundscape Part 2: Data Collection and Reporting Requirements*; ISO: Geneva, Switzerland, 2018.
35. Carles, J.; Bernáldez, F.; Lucio, J.D. Audio-visual interactions and soundscape preferences. *Landsc. Res.* **1992**, *17*, 52–56. [[CrossRef](#)]
36. Galbrun, L.; Calarco, F.M. Audio-visual interaction and perceptual assessment of water features used over road traffic noise. *J. Acoust. Soc. Am.* **2014**, *136*, 2609–2620. [[CrossRef](#)]



37. Hong, J.Y.; Jeon, J.Y. The effects of audio–visual factors on perceptions of environmental noise barrier performance. *Landsc. Urban Plan.* **2014**, *125*, 28–37. [[CrossRef](#)]
38. Liu, F.; Kang, J. Relationship between street scale and subjective assessment of audio-visual environment comfort based on 3D virtual reality and dual-channel acoustic tests. *Build. Environ.* **2018**, *129*, 35–45. [[CrossRef](#)]
39. Preis, A.; Kociński, J.; Hafke-Dys, H.; Wrzosek, M. Audio-visual interactions in environment assessment. *Sci. Total. Environ.* **2015**, *523*, 191–200. [[CrossRef](#)]
40. Szychowska, M.; Hafke-Dys, H.; Preis, A.; Kociński, J.; Kleka, P. The influence of audio-visual interactions on the annoyance ratings for wind turbines. *Appl. Acoust.* **2018**, *129*, 190–203. [[CrossRef](#)]
41. Viollon, S.; Lavandier, C.; Drake, C. Influence of visual setting on sound ratings in an urban environment. *Appl. Acoust.* **2002**, *63*, 493–511. [[CrossRef](#)]
42. Muhammad, I.; Vorländer, M.; Schlittmeier, S.J. Audio-video virtual reality environments in building acoustics: An exemplary study reproducing performance results and subjective ratings of a laboratory listening experiment. *J. Acoust. Soc. Am.* **2019**, *146*, EL310–EL316. [[CrossRef](#)]
43. Thery, D.; Boccarda, V.; Katz, B.F. Auralization uses in acoustical design: A survey study of acoustical consultants. *J. Acoust. Soc. Am.* **2019**, *145*, 3446–3456. [[CrossRef](#)]
44. Postma, B.N.; Katz, B.F. The influence of visual distance on the room-acoustic experience of auralizations. *J. Acoust. Soc. Am.* **2017**, *142*, 3035–3046. [[CrossRef](#)]
45. Abdalrahman, Z.; Galbrun, L. Audio-visual preferences, perception, and use of water features in open-plan offices. *J. Acoust. Soc. Am.* **2020**, *147*, 1661–1672. [[CrossRef](#)]
46. Picinali, L.; Afonso, A.; Denis, M.; Katz, B.F. Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge. *Int. J. Hum. Comput. Stud.* **2014**, *72*, 393–407. [[CrossRef](#)]
47. Afonso, A.; Blum, A.; Katz, B.F.; Tarroux, P.; Borst, G.; Denis, M. Structural properties of spatial representations in blind people: Scanning images constructed from haptic exploration or from locomotion in a 3-D audio virtual environment. *Mem. Cogn.* **2010**, *38*, 591–604. [[CrossRef](#)]
48. Brown, A.; Kang, J.; Gjestland, T. Towards standardization in soundscape preference assessment. *Appl. Acoust.* **2011**, *72*, 387–392. [[CrossRef](#)]
49. Kang, J.; Yu, L. Modelling subjective evaluation of soundscape: Towards soundscape standardization. *J. Acoust. Soc. Am.* **2011**, *129*, 2570. [[CrossRef](#)]
50. ISO. *TS 12913-3: 2018—Acoustics—Soundscape Part 3: Data Analysis*; ISO: Geneva, Switzerland, 2019.
51. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.C.; Ioannidis, J.P.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *J. Clin. Epidemiol.* **2009**, *62*, e1–e34. [[CrossRef](#)]
52. Jeon, J.Y.; Jo, H.I. Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment. *Build. Environ.* **2020**, *169*, 106544. [[CrossRef](#)]
53. Aletta, F.; Masullo, M.; Maffei, L.; Kang, J. The effect of vision on the perception of the noise produced by a chiller in a common living environment. *Noise Control. Eng. J.* **2016**, *64*, 363–378. [[CrossRef](#)]
54. Iachini, T.; Maffei, L.; Ruotolo, F.; Senese, V.P.; Ruggiero, G.; Masullo, M.; Alekseeva, N. Multisensory assessment of acoustic comfort aboard metros: A virtual reality study. *Appl. Cogn. Psychol.* **2012**, *26*, 757–767. [[CrossRef](#)]
55. Maffei, L.; Iachini, T.; Masullo, M.; Aletta, F.; Sorrentino, F.; Senese, V.P.; Ruotolo, F. The effects of vision-related aspects on noise perception of wind turbines in quiet areas. *Int. J. Environ. Res. Public Health* **2013**, *10*, 1681–1697. [[CrossRef](#)]
56. Maffei, L.; Masullo, M.; Aletta, F.; Di Gabriele, M. The influence of visual characteristics of barriers on railway noise perception. *Sci. Total Environ.* **2013**, *445*, 41–47. [[CrossRef](#)]
57. Sanchez, G.M.E.; Van Renterghem, T.; Sun, K.; De Coensel, B.; Botteldooren, D. Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. *Landsc. Urban Plan.* **2017**, *167*, 98–107. [[CrossRef](#)]
58. Yu, T.; Behm, H.; Bill, R.; Kang, J. Audio-visual perception of new wind parks. *Landsc. Urban Plan.* **2017**, *165*, 1–10. [[CrossRef](#)]
59. Hong, J.Y.; Lam, B.; Ong, Z.-T.; Ooi, K.; Gan, W.-S.; Kang, J.; Feng, J.; Tan, S.-T. Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. *Build. Environ.* **2019**, *149*, 1–14. [[CrossRef](#)]
60. Jeon, J.Y.; Jo, H.I. Three-dimensional virtual reality-based subjective evaluation of road traffic noise heard in urban high-rise residential buildings. *Build. Environ.* **2019**, *148*, 468–477. [[CrossRef](#)]
61. Sun, K.; De Coensel, B.; Filipan, K.; Aletta, F.; Van Renterghem, T.; De Pessemer, T.; Joseph, W.; Botteldooren, D. Classification of soundscapes of urban public open spaces. *Landsc. Urban Plan.* **2019**, *189*, 139–155. [[CrossRef](#)]
62. Kang, J. *Urban Sound Environment*; CRC Press: Boca Raton, FL, USA, 2006. [[CrossRef](#)]
63. Basner, M.; Babisch, W.; Davis, A.; Brink, M.; Clark, C.; Janssen, S.; Stansfeld, S. Auditory and non-auditory effects of noise on health. *Lancet* **2014**, *383*, 1325–1332. [[CrossRef](#)]
64. Stansfeld, S.A.; Matheson, M.P. Noise pollution: Non-auditory effects on health. *Br. Med. Bull.* **2003**, *68*, 243–257. [[CrossRef](#)]
65. Annerstedt, M.; Jönsson, P.; Wallergård, M.; Johansson, G.; Karlsson, B.; Grahn, P.; Hansen, Å.M.; Währborg, P. Inducing physiological stress recovery with sounds of nature in a virtual reality forest—Results from a pilot study. *Physiol. Behav.* **2013**, *118*, 240–250. [[CrossRef](#)]

66. Hedblom, M.; Gunnarsson, B.; Schaefer, M.; Knez, I.; Thorsson, P.; Lundström, J.N. Sounds of nature in the city: No evidence of bird song improving stress recovery. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1390. [[CrossRef](#)]
67. Jiang, L.; Masullo, M.; Maffei, L.; Meng, F.; Vorländer, M. How do shared-street design and traffic restriction improve urban soundscape and human experience?—An online survey with virtual reality. *Build. Environ.* **2018**, *143*, 318–328. [[CrossRef](#)]
68. Jiang, L.; Masullo, M.; Maffei, L.; Meng, F.; Vorländer, M. A demonstrator tool of web-based virtual reality for participatory evaluation of urban sound environment. *Landsc. Urban Plan.* **2018**, *170*, 276–282. [[CrossRef](#)]
69. Jahncke, H.; Eriksson, K.; Naula, S. The effects of auditive and visual settings on perceived restoration likelihood. *Noise Health* **2015**, *17*, 1. [[CrossRef](#)]
70. Guastavino, C.; Katz, B.F. Perceptual evaluation of multi-dimensional spatial audio reproduction. *J. Acoust. Soc. Am.* **2004**, *116*, 1105–1115. [[CrossRef](#)]
71. Jiang, L.; Kang, J. Effect of traffic noise on perceived visual impact of motorway traffic. *Landsc. Urban Plan.* **2016**, *150*, 50–59. [[CrossRef](#)]
72. Chau, C.K.; Leung, T.M.; Xu, J.M.; Tang, S.K. Modelling noise annoyance responses to combined sound sources and views of sea, road traffic, and mountain greenery. *J. Acoust. Soc. Am.* **2018**, *144*, 3503–3513. [[CrossRef](#)] [[PubMed](#)]
73. Thery, D.; Poirier-Quinot, D.; Postma, B.N.; Katz, B.F. Impact of The Visual Rendering System on Subjective Auralization Assessment in VR. In *Proceedings of the International Conference on Virtual Reality and Augmented Reality*; Springer: Berlin, Germany, 2017; pp. 105–118.
74. Cruz-Neira, C.; Sandin, D.J.; DeFanti, T.A.; Kenyon, R.V.; Hart, J.C. The CAVE: Audio visual experience automatic virtual environment. *Commun. ACM* **1992**, *35*, 64–73. [[CrossRef](#)]
75. Muhanna, M.A. Virtual reality and the CAVE: Taxonomy, interaction challenges and research directions. *J. King Saud Univ. Comput. Inf. Sci.* **2015**, *27*, 344–361. [[CrossRef](#)]
76. Havig, P.; McIntire, J.; Geiselman, E. *Virtual Reality in a Cave: Limitations and the Need for HMDs? Head-and Helmet-Mounted Displays XVI: Design and Applications*; International Society for Optics and Photonics: Bellingham, WA, USA, 1 June 2011; p. 804107.
77. Postma, B.N.; Katz, B.F. Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *J. Acoust. Soc. Am.* **2016**, *140*, 4326–4337. [[CrossRef](#)]
78. Lokki, T.; McLeod, L.; Kuusinen, A. Perception of loudness and envelopment for different orchestral dynamics. *J. Acoust. Soc. Am.* **2020**, *148*, 2137–2145. [[CrossRef](#)]
79. Poirier-Quinot, D.; Postma, B.N.; Katz, B.F. Augmented Auralization: Complimenting Auralizations with Immersive Virtual Reality Technologies. In *Proceedings of the International Symposium on Music and Room Acoustics (ISMRA)*, La Plata, Argentina, 11–13 September 2016; pp. 1–10.
80. Schissler, C.; Nicholls, A.; Mehra, R. Efficient HRTF-based spatial audio for area and volumetric sources. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 1356–1366. [[CrossRef](#)] [[PubMed](#)]
81. Serafin, S.; Geronazzo, M.; Erkut, C.; Nilsson, N.C.; Nordahl, R. Sonic interactions in virtual reality: State of the art, current challenges, and future directions. *IEEE Comput. Graph. Appl.* **2018**, *38*, 31–43. [[CrossRef](#)]
82. Xu, C.; Kang, J. Soundscape evaluation: Binaural or monaural? *J. Acoust. Soc. Am.* **2019**, *145*, 3208–3217. [[CrossRef](#)]
83. Oberman, T.; Jambrošić, K.; Aletta, F.; Kang, J. Towards a Soundscape Surround Index. In *Proceedings of the International Conference on Acoustics*, Aachen, Germany, 9–13 September 2019; pp. 9–13.
84. Narbutt, M.; O’Leary, S.; Allen, A.; Skoglund, J.; Hines, A. Streaming VR for immersion: Quality aspects of compressed spatial audio. In *Proceedings of the 2017 23rd International Conference on Virtual System & Multimedia (VSMM)*; IEEE: Dublin, Ireland, 2017; pp. 1–6.
85. Pausch, F.; Aspöck, L.; Vorländer, M.; Fels, J. An extended binaural real-time auralization system with an interface to research hearing aids for experiments on subjects with hearing loss. *Trends Hear.* **2018**, *22*. [[CrossRef](#)]
86. Favrot, S.; Buchholz, J.M. LoRA: A loudspeaker-based room auralization system. *Acta Acust. United Acust.* **2010**, *96*, 364–375. [[CrossRef](#)]
87. Perotin, L.; Serizel, R.; Vincent, E.; Guérin, A. Multichannel Speech Separation with Recurrent Neural Networks from High-Order Ambisonics Recordings. In *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)*, Calgary, AB, Canada, 15–20 April 2018; pp. 36–40.