

Article

Assessing Ambisonics Sound Source Localization by Means of Virtual Reality and Gamification Tools

Esaú Medina ¹, Rhoddy Viveros-Muñoz ² and Felipe Otondo ^{1,*}

¹ Arts and Technology Lab, Institute of Acoustics, Universidad Austral, Valdivia 5110701, Chile; esau.medina@alumnos.uach.cl

² Departamento de Electrónica e Informática, Universidad Técnica Federico Santa María, Sede Concepción, Hualpén 4601086, Chile; rhoddy.viveros@usm.cl

* Correspondence: felipe.otondo@uach.cl

Abstract: Sound localization is a key area of interest in auditory research, especially in complex acoustic environments. This study evaluates the impact of incorporating higher-order Ambisonics (HOA) with virtual reality (VR) and gamification tools on sound source localization. The research addresses the current limitations in VR audio systems, particularly the lack of native support for HOA in game engines like Unreal Engine (UE). A novel framework was developed, combining UE for VR graphics rendering and Max for HOA audio processing. Participants performed sound source localization tasks in two VR environments using a head-mounted display (HMD). The assessment included both horizontal and vertical plane localization. Gamification elements were introduced to improve engagement and task comprehension. Results showed significant improvements in horizontal localization accuracy, although challenges remained in back localization. The findings underscore the potential of VR and gamification to enhance auditory tests, reducing test duration and participant fatigue. This research contributes to the development of immersive and interactive audio experiences, highlighting the broader applications of VR beyond entertainment.

Keywords: sound perception; Ambisonics; virtual reality; gamification; 3D audio



Citation: Medina, E.; Viveros-Muñoz, R.; Otondo, F. Assessing Ambisonics Sound Source Localization by Means of Virtual Reality and Gamification Tools. *Appl. Sci.* **2024**, *14*, 7986. <https://doi.org/10.3390/app14177986>

Academic Editor: Gino Iannace

Received: 30 July 2024

Revised: 22 August 2024

Accepted: 2 September 2024

Published: 6 September 2024



Copyright: © 2024 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

In recent years, virtual reality (VR) has surged in popularity and accessibility, extending its reach beyond gaming and becoming integral to diverse fields such as scientific research, healthcare, education, and the arts. These developments have also extended into the domain of auditory research, where VR has proven to be a versatile tool for training tasks, perceptual evaluation, and immersive audio development [1–3].

An essential aspect of VR audio research lies in the reproduction of auditory stimuli as well as the evaluation of the subject's response to the audio playback system [4]. The use of headphones has often been preferred for practical reasons, such as sound control and isolation, reduced sound reflections, the option for binaural tests, and the use of generic or native head-related transfer functions (HRTFs) [5–7]. Nevertheless, recent studies, particularly those focusing on sound source localization, opt for loudspeaker reproduction [8,9]. This choice allows for a faithful simulation of the real environment, preserving interaural cues, far-field conditions, and other important auditory factors. In this regard, sound field synthesis (SFS) methods such as Ambisonics have emerged as a suitable alternative for loudspeaker array playback [10].

Ambisonics, a sound encoding method, generates a sound field from diverse loudspeaker array setups, offering flexibility and precision in 3D audio reproduction [11]. It achieves this by encoding the sound field with a larger number of audio channels. This increased resolution enables higher-order Ambisonics (HOA) to better preserve the nuances of the sound field, including the direction and distance of sound sources, and to recreate a more accurate and engaging 3D audio environment. Traditional Ambisonics reproduction

uses four channels (1st order), whereas HOA typically employs nine channels (2nd order) or higher, using advanced Ambisonics encoders, which hold promise for improved audio reproduction. It becomes interesting, then, to analyze the possibility of improving the perception of sound localization in VR using HOA playback. However, achieving accurate localization in an HOA setup requires an increase in the number of available channels to match theoretical requirements.

Recent studies related to immersive audio VR systems have explored the influence of visual cues on error angle reduction, reaction time, and compensation for sound cue alterations induced by head-mounted displays (HMDs) in both VR-based and gamified sound source localization tests [9,12]. However, despite these investigations, there remains a gap in comprehensive research examining the integration of VR graphics and gamification elements in such tests. Furthermore, the lack of native compatibility between existing game engines like Unreal Engine (UE) and HOA encoding presents a significant obstacle for researchers and developers alike, thus limiting the integration of this audio encoding into virtual environments [13]. To address this issue, a novel framework combining UE for VR graphics rendering and Max for HOA audio processing was developed [14]. This builds upon previous pilot projects indicating the reliability of HOA encoding for sound source spatialization in specific room characteristics [15]. Thus, this study explores the impact of integrating VR graphics and gamification on sound source localization within virtual environments. Specifically, it aims to address gaps in understanding their effects on sound source localization. This leads to the research question: “Does integrating VR environments and gamification tools enhance subjects’ performance in 3D audio localization listening tests?” It was hypothesized that this integration would improve subjects’ performance, facilitated by enhanced task comprehension and visual cues that compensate for HMD-induced alterations in sound perception.

The experimental design involved participants performing a sound source localization task in a VR version of the acoustic laboratory where the experiments were conducted, utilizing 2nd-order HOA reproduction for auditory stimulus playback. Gamification elements such as buttons, animations, interactive instructions, and a progress bar were integrated into the VR environment to enhance task comprehension.

2. Related Work

2.1. Sound Source Localization in Ambisonics Systems

Virtual sound source localization tasks have been widely used for measuring the performance of spatial audio systems and auditory research [16,17]. Ambisonics is often preferred by researchers for experiments involving both users’ natural auditory conditions with loudspeakers and binaural reproduction using HRTF with headphones. P. Power et al. conducted an evaluation of sound source localization accuracy across various elevation and azimuth positions [18], comparing different Ambisonics orders using a loudspeaker array. This study focused on the performance of Ambisonics systems in accurately localizing virtual sound sources in the vertical plane. K. Sochaczewska assessed the spatial resolution of a 3rd-order Ambisonics system by conducting a listening test in which participants identified the direction of sound sources [19]. This approach allowed for an evaluation of the system’s perceived spatial accuracy, providing insights into the effectiveness of Ambisonics-encoded sound sources. T. Rudzki et al. investigated the influence of Ambisonics order and encoding bitrate on the localization of virtual sound sources reproduced over both loudspeaker arrays and binaural systems [20]. Similarly, L. Thresh et al. assessed the performance of multiple-order Ambisonics systems through a sound source localization test comparing real and virtual sound sources [21]. In these works, it was highlighted that the Ambisonics order had a direct effect on sources’ localization, with HOA showing a clear improvement over the 1st-order setup. Another key finding was that this trend is more prominent when conducting loudspeaker-based tests.

2.2. Virtual Reality and Auditory Research

Some recent studies related to auditory research, particularly to sound source localization, have been performed in addition to VR environments. A study by C. Valzolgher investigated the impact of the presence of visual frames during a sound source localization task, finding that sound source localization improved when visual elements were present even without explicit visual cues about the stimuli [12]. A sound source localization experiment in an anechoic chamber was conducted using an Ambisonics setup in a virtual test room by T. Huisman et al. [8]. They assessed the effects under three conditions: blindfolded, in the VR room without visible loudspeakers, and in the VR room with visible loudspeakers. Their key finding was that increasing the number of visual references related to the sound sources improved the performance in terms of sound source localization of the participants and compensated for the effects of using an HMD.

A training application focused on improving sound localization with binaural playback was designed by Y. Wu, taking advantage of VR development [6]. Results of the study showed a decrease in the overall error and reaction times among participants. This capturing system of users' performance provides an alternative for gathering data from a framework built from a video game engine. M. Frank et al. recreated a multiple loudspeaker experiment to evaluate an HOA hemisphere array in VR [7]. To achieve this, the authors used open-source VST for binaural audio rendering along with Unity as the graphic engine. Similarly, M. Piotrowska et al. developed a system that integrates Ambisonics spatial audio within a virtual environment for conducting sound source localization tests [22]. They highlighted the advantages of using VR technology in auditory research, including precise control and flexibility in experimental conditions, achieved through the use of audio software in conjunction with the game engine.

A gamified localization task was developed by J. Rees-Jones with the implementation of a hybrid platform combining the game engine and specialized audio software such as Max for sound source spatialization on different loudspeaker systems [9]. The gamified approach in this experiment was intentionally simplified to avoid distractions from the auditory task. However, an interest in conducting a similar experiment with a more complex task and environment design was declared by the authors. M. Ramirez evaluated sound localization accuracy using conventional loudspeaker setups and binaural presentation with nonindividual HRTF alongside various VR-based environments [23]. The study found that virtualization of both auditory stimuli and visual information increased localization blur but also demonstrated that VR testing could be a practical and efficient alternative to traditional methods. The results of these studies support the development of VR applications for assessing spatial hearing abilities and suggest that VR could advance research in spatial auditory processing by providing a time- and resource-efficient approach to exploring auditory perception.

2.3. The Soundlapse Project and 3D Audio Research

The research presented here is part of the activities of the Soundlapse project, which explores the acoustical features of natural soundscapes in southern Chile [24]. Among the project's research lines, there is a focus on multichannel audio reproduction and its applications in diverse contexts such as arts, education, and VR. Previous studies on this topic have emphasized the need for accurate 3D Ambisonics reproduction across various applications. These studies included designing and implementing listening tests to assess the quality of Ambisonics in sound localization and immersion [15,25].

The first attempt to deal with this issue was carried out by J. Fernández, who investigated a spatial audio design over a binaural system by conducting a sound localization test [25]. An original algorithm was developed using the software Max 8 and *Spat* libraries [26]. Following this study, M. Aguilar assessed the performance of an optimized version of the Ambisonics algorithm by conducting a localization test in different architectural conditions [15]. A key finding of this study was that substantial changes in the acoustic attributes of the rooms where the experiments took place had a minimal impact on

participants' sound source localization performance. Results also showed the importance of the flexibility of the designed Max tool for research and artistic purposes and the limitations of the system for sound source spatialization, particularly in the vertical plane.

3. Materials and Methods

3.1. Participants

Thirty-one undergraduate and postgraduate students from the Universidad Austral in Valdivia, Chile, took part in the study. Participants were aged between 18 and 35 years, including 11 women and 20 men. All reported normal hearing abilities and no prior experience with the experimental tasks. Participants were recruited through advertisements posted on the university's social media and provided informed consent before taking part in the study. Participants were free to interrupt the experiment at any time if needed. Participation in the study was entirely voluntary, with no financial compensation offered.

3.2. Experimental Setup

The tests took place at the Arts and Technology Lab (LATE) at Universidad Austral [27]. The experimental setup involved an original VR 3D audio environment designed in UE for graphics rendering and Max software with *Spat* libraries for higher-order Ambisonics (HOA) audio processing.

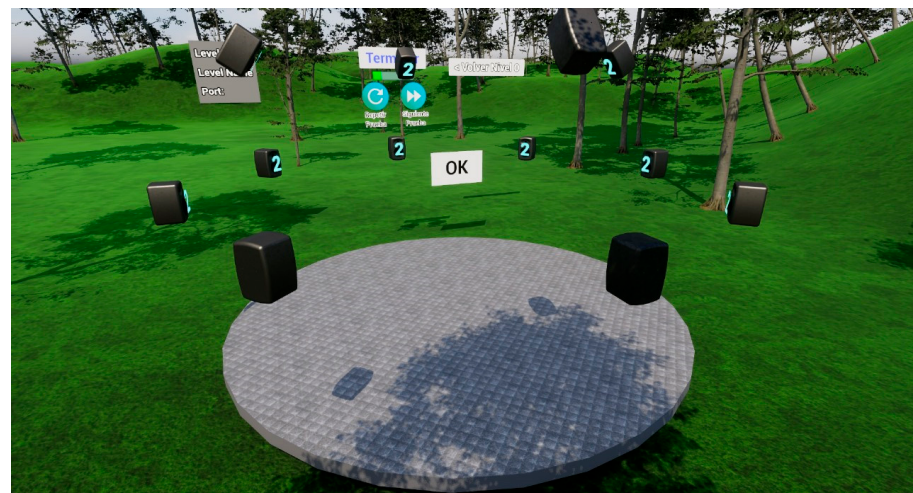
During the task, participants wore an HMD and were asked to determine the direction from which an auditory stimulus originates. The experiment was conducted using 12 Genelec 8330 loudspeakers (Genelec, Iisalmi, Finland) arranged in a hemispheric array configured for Ambisonics reproduction. Objective performance measures, including accuracy of sound localization, were recorded during the task. Additionally, subjective measures such as participant feedback and perceived task difficulty were collected through a post-task informal interview.

The VR environment was created using UE's Blueprints visual scripting language (Unreal Engine version 5.3), enabling communication via OSC with a Max patch for stimulus playback. In the first part of the test, the VR environment was designed to replicate the LATE lab (Figure 1a), featuring accurate spatial representations of the Genelec 8330 loudspeakers arranged according to the virtual source distribution shown in Figure 2.



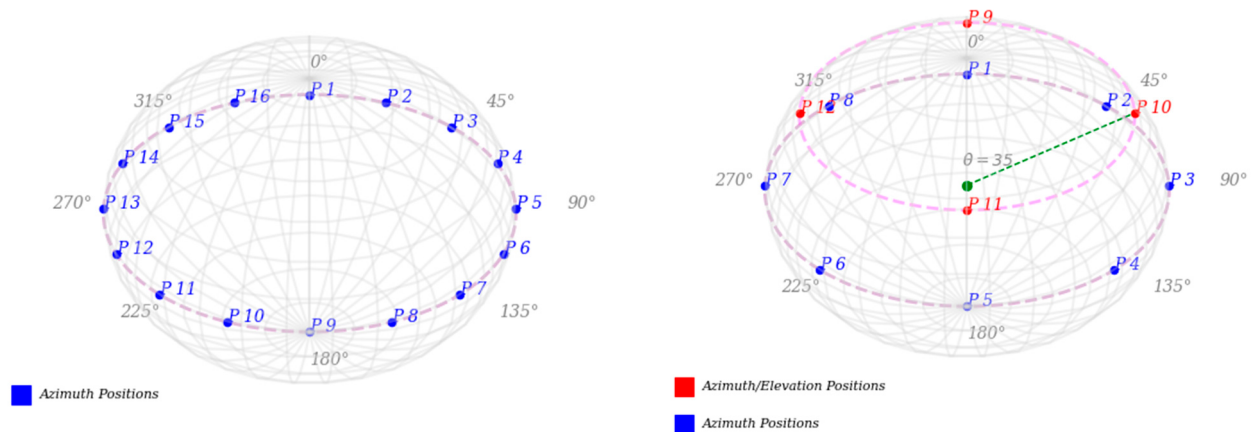
(a)

Figure 1. Cont.



(b)

Figure 1. (a) VR environment as a replica of the LATE lab at Universidad Austral de Chile. (b) VR environment of a wetland in the city of Valdivia, Chile.



(a)

(b)

Figure 2. Spatial representations of the loudspeakers, arranged according to the virtual source distribution. (a) Horizontal distribution at 0° elevation. (b) Horizontal distribution at 0° and 35° elevation.

In the second part of the test, a wetland 2nd-order Ambisonics field recording was played in the background to enhance the immersive experience for listeners and add realism to the virtual environment as shown in Figure 1b.

The designed virtual environments integrated various gamification elements to improve task engagement and comprehension among participants. These features included interactive buttons for controlling stimulus playback, a progress bar to indicate task progress, and VR avatar elements like virtual hands for interacting with the environment, as shown in Figure 3.

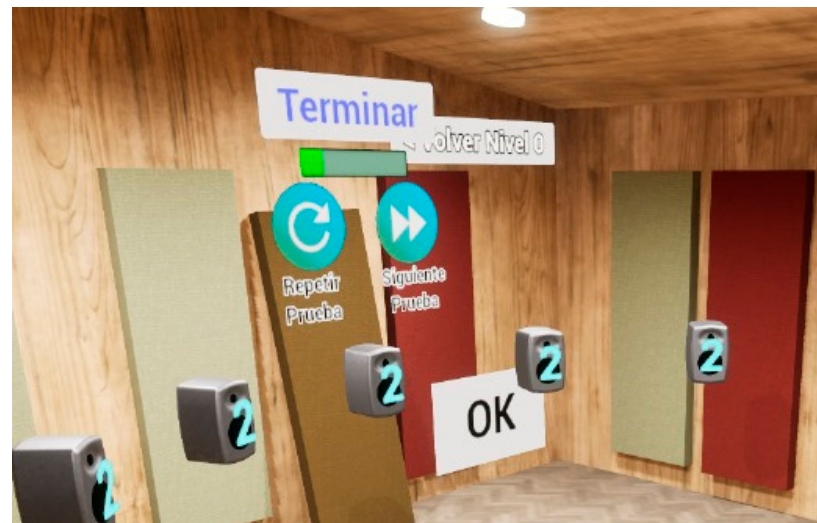


Figure 3. User interface for stimulus control within the VR environment.

3.3. Auditory Stimuli

The auditory stimuli used in this study were the call of the Southern Lapwing (*Vanellus chilensis*), commonly referred to as the ‘Queltehue’. This choice was influenced by its relevance to the *Soundlapse* project, which explores the natural soundscape of Southern Chile. The Queltehue’s call was selected for its broad spectral characteristics, especially its content around 4 kHz and higher, which are known to significantly enhance vertical and front/back localization [28]. The sound file, extracted from a soundscape recording, was also used in previous research by J. Fernández and M. Aguilar [12,16] (Figure 4). It was in mono format and spatialized as a virtual sound source using a 2nd-order Ambisonics encoder in an original patch created in Max with the Spat libraries (Figure 5).

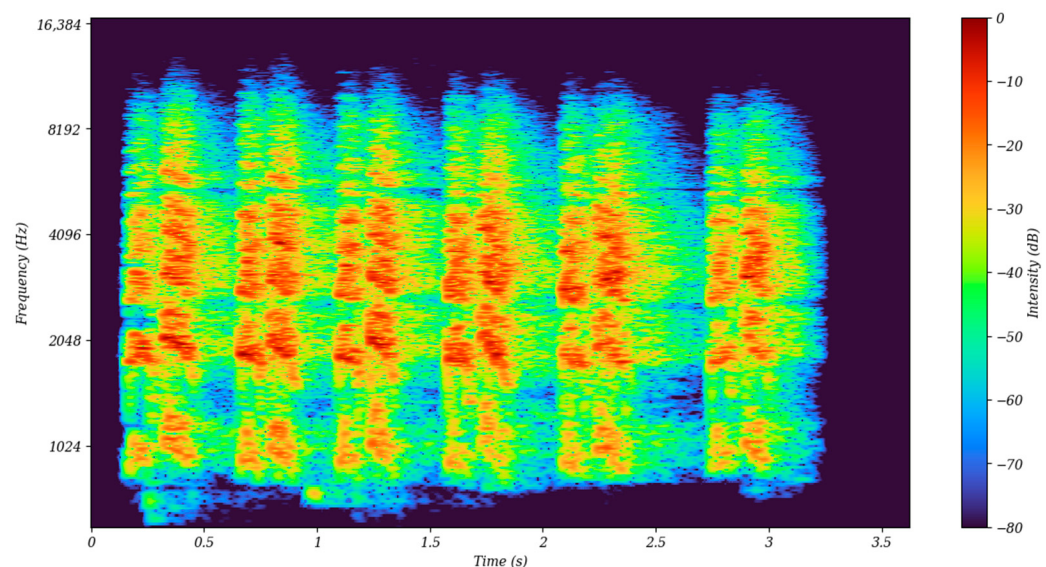


Figure 4. Spectrogram of the auditory stimulus used in this study: call of the bird *Vanellus chilensis*, showing frequency content of the call, emphasizing the spectral range with significant components relevant for vertical and front/back localization.

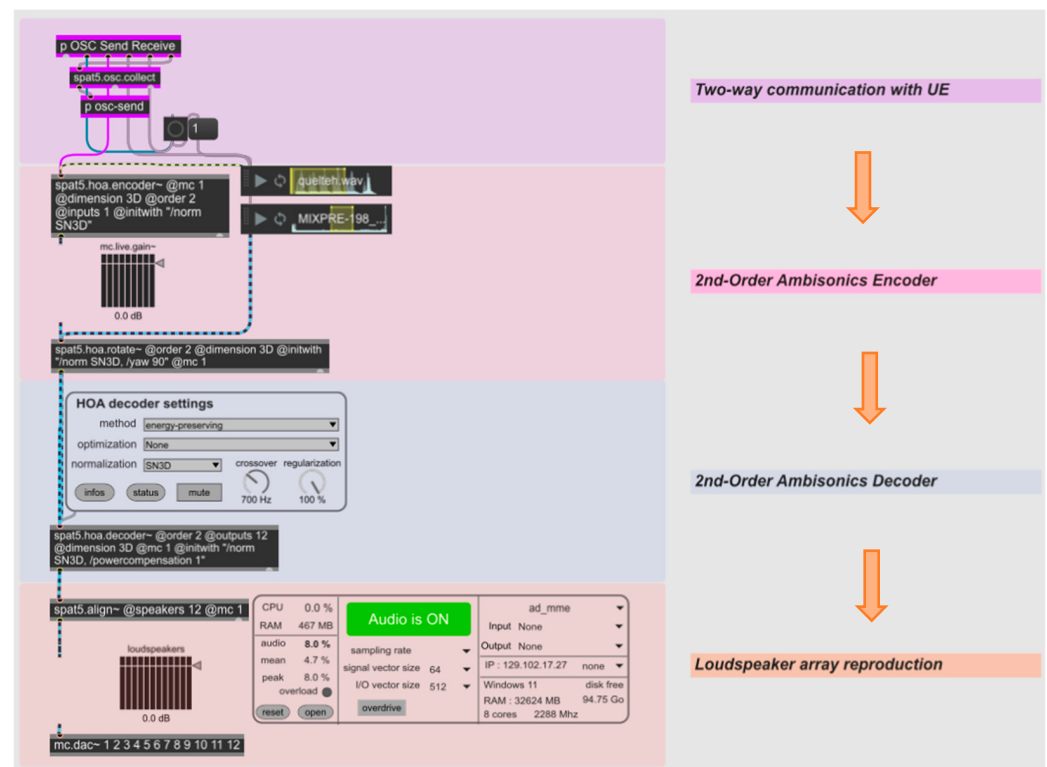


Figure 5. Flow in Max/*Spat* used for sound localization experiments with two-way OSC communication with UE.

In the first part of the experiment, stimuli were randomly presented from 16 virtual sound sources at specific positions on the horizontal plane (Figure 2a). In the second part of the test, stimuli were presented from 12 positions, including some with an elevation angle of 35° to introduce vertical localization tasks (Figure 2b).

The reproduced sound pressure level of the stimuli was calibrated to ensure consistent loudness across all trials. The randomization of stimulus presentation positions in both parts of the experiment helped prevent learning effects and ensured that participants had to rely on their auditory localization abilities.

3.4. Experimental Task

The experimental task was divided into two parts. In the first round, designed to assess sound localization in the horizontal plane, the auditory stimulus was presented randomly from 16 positions distributed equidistantly every 22.5° around the listener. Each position was repeated twice, resulting in a total of 32 trials for the first round. This first round was conducted in the virtual replica of the LATe laboratory environment mentioned above and shown in Figures 1a and 2a.

In the second round, the task was performed in a virtual wetland environment with the stimulus presented from 12 different positions. Eight positions were distributed every 45° in the horizontal plane, and four additional positions were placed every 90° at a vertical elevation of 35° (Figure 2b). This new arrangement was designed to integrate both horizontal and vertical localization tests.

For both rounds, the stimuli were presented randomly when the participant pressed the play button within the VR environment. Participants were instructed to face forward before each stimulus was played to maintain a consistent reference point. After listening to the sound, they could turn or rotate to indicate the perceived position using a VR controller. However, they were required to return to the forward-facing position before playing the next stimulus. This approach minimized variability, allowing for fairer comparisons across trials. Fixing their head position during playback prevented additional variables

from affecting sound localization, addressing the challenge of locating sounds when free movement is allowed.

Participants' responses were continually saved and stored in a file generated by the Max patch during the performance of the test.

3.5. Data Analysis

Data were collected using the Max patch, which stored pairs of data points showing the original virtual sound source positions and the perceived position for each participant. The task was divided into two rounds, resulting in a comprehensive dataset for each participant that included responses from both the horizontal plane localization test and the elevated plane localization test. The resulting data statistical analysis was performed using Python 3.12, with libraries such as NumPy, Pandas, Seaborn, and Matplotlib. Circular statistics were applied to assess the angular characteristics of the data [29–34].

4. Results

4.1. Sound Localization First Round: Horizontal Plane

The distribution of angular localization perception for the horizontal plane positions in the first round is shown in the violin plot below (Figure 6). Because the data are angular positions at a single fixed azimuth, we considered using circular statistics to address the results.

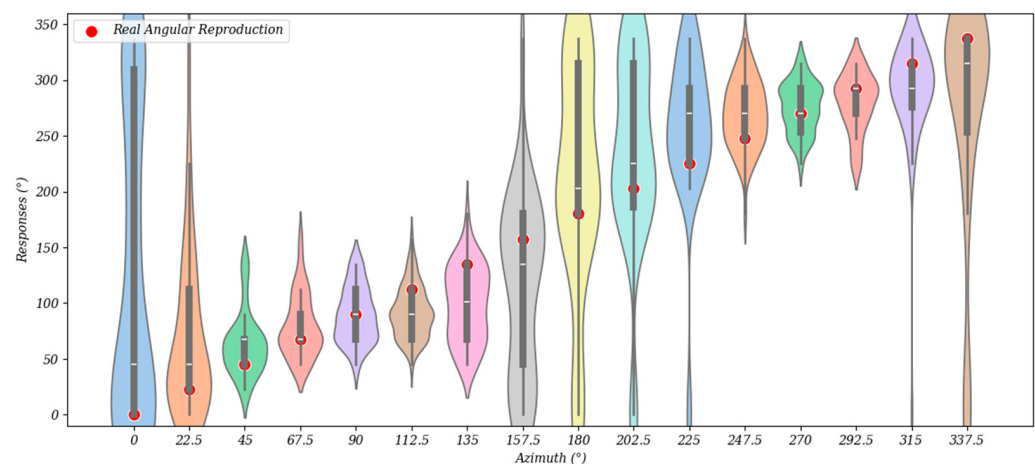


Figure 6. Violin plot of sound localization perception in the first round. The distribution of localization responses is shown as a function of the azimuth. The violin plot illustrates the spread of responses per angle, with the shape representing a sideways histogram.

To determine whether the participants' perception of location was correctly aligned with the position where the audio was played, we employed the nonparametric V-test [29,34]. The V-test was used to assess whether there was a significant difference between the mean response angle and the expected angle for each position, with the null hypothesis assuming a uniform distribution of the responses.

Results indicate that the angular differences were statistically significant at most positions, allowing us to reject the null hypothesis at an alpha level of 0.05. However, at the position of 180°, the difference in the mean angle was notably distinct, highlighting a significant deviation in localization accuracy at this specific point. The results of the V-test for each position, including the mean angle and the V and u statistics, are presented in Table 1, with significance determined at an alpha level of 0.05.

Table 1. V-test results for the first round (horizontal plane).

Real Position (deg)	Perception Mean Angle (deg)	V	u	Significant ($\alpha = 0.05$)
0	354.05	37.38	6.71	TRUE
22.5	44.60	33.85	6.08	TRUE
45	61.77	52.60	9.44	TRUE
67.5	78.11	54.16	9.72	TRUE
90	88.98	56.48	10.14	TRUE
112.5	91.83	53.61	9.62	TRUE
135	99.00	41.99	7.54	TRUE
157.5	116.71	36.73	3.00	TRUE
180	253.92	4.585	0.82	FALSE
202.5	255.34	16.60	2.98	TRUE
225	269.93	31.58	5.67	TRUE
247.5	250.21	50.21	9.01	TRUE
270	275.56	57.13	10.2	TRUE
292.5	280.41	54.63	9.81	TRUE
315	297.85	51.93	9.32	TRUE
337.5	321.94	44.76	8.04	TRUE

4.2. Sound Localization Second Round: Horizontal and Vertical Planes

In the second round, sound localization in the horizontal and vertical planes was assessed, generating responses with both azimuth and elevation information. Consequently, the results of this second round are treated as spherical data. First, it was necessary to evaluate the normality of the data to determine whether to use parametric or non-parametric statistics. The Rayleigh test was employed to assess the normality of the data [31].

The Rayleigh test statistic (Z) was 115.04 ($p < 0.001$), indicating that the sound localizations were significantly different from a uniform distribution ($\alpha = 0.05$). This suggests that the data are clustered or concentrated in specific directions. Therefore, the next step was to analyze whether these groupings differed from the actual positions where the stimuli were played.

Given that the data do not conform to normality, a non-parametric statistical analysis was necessary to assess whether the perceived locations matched the actual locations where the stimuli were played. The Hodges–Ajne test provides a non-parametric measure to compare a fixed angular position, which, in this case, is the actual location where the audio was played, against the population mean of participants' responses [32,33]. The results of the Hodges–Ajne test are shown in Figure 7. Significant differences were observed for the angles 0° , 45° , 90° , 135° , 225° , 270° , and 315° at both elevations ($p < 0.05$), indicating that the sound localizations were significantly different from a uniform distribution but aligned with the actual location where the stimuli were played. In contrast, for the angles 180° at both elevations (0° and 35°), the p -values ($p = 0.121$) suggest that the responses were not significantly different from a uniform distribution around the reproduced angle.

It is also interesting to analyze whether the participants' responses were different in perception of elevation. That case only occurs at four angles where the sound was reproduced at 0° and 35° elevation. These cases are at 0° , 90° , 180° , and 270° . In this case, we want to compare two population samples—responses for elevation at 0° and at 35° —at the four angles evaluated (0° , 90° , 180° , and 270°).

The Mardia–Watson–Wheeler (MWW) test is a nonparametric statistical method used to compare the distributions of two or more angular populations [30]. Therefore, the MWW test was performed for the angles 0° , 90° , 180° , and 270° at two different elevations (0° and 35°). The results indicate a significant main effect of elevation on the sound localization data across the angles 0° , 90° , 180° , and 270° ($\chi^2 = 26.9877$, $p < 0.0001$), suggesting that the distributions of responses differ significantly between the two elevations.

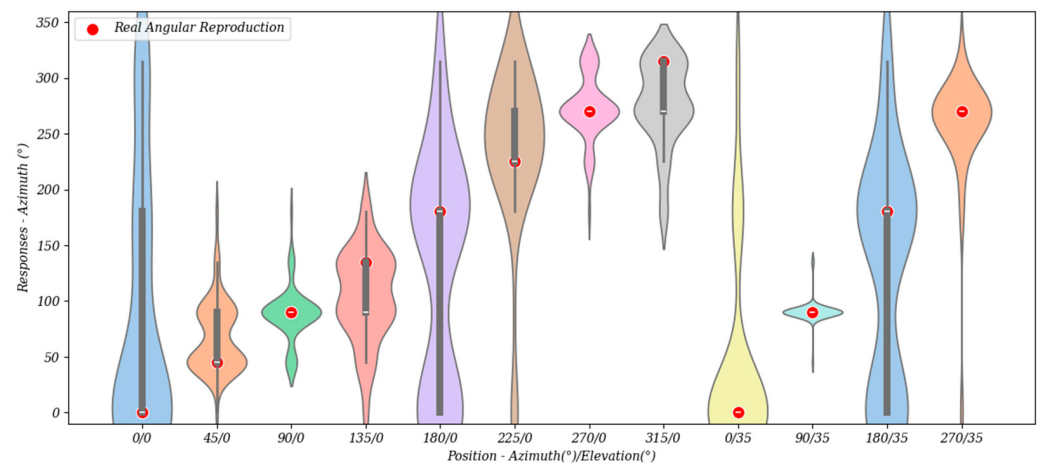


Figure 7. Violin plot of sound localization perception in the second round. The distribution of localization responses is shown as a function of the azimuth. It presents the cases with both elevations of 0° and 35°. The violin plot illustrates the spread of responses per angle, with the shape representing a sideways histogram.

Post hoc comparisons were conducted to further investigate the differences between the elevations for each specific angle. The results are summarized in Table 2.

Table 2. MWW test results for the second round.

Comparison	Chi-Squared Statistic	<i>p</i> -Value
0° elevation 0° vs. 35°	1.9283	0.1649
90° elevation 0° vs. 35°	13.7125	0.0002
180° elevation 0° vs. 35°	0.8570	0.3546
270° elevation 0° vs. 35°	8.5941	0.0034

Significant differences were found between the elevations at 90° ($\chi^2 = 13.7125$, $p < 0.001$) and 270° ($\chi^2 = 8.5941$, $p < 0.01$), indicating that the responses were significantly different between elevations 0° and 35° at these angles. However, no significant differences were observed for the angles 0° ($\chi^2 = 1.9283$, $p = 0.1649$) and 180° ($\chi^2 = 0.8570$, $p = 0.3546$), suggesting that the responses were not significantly different between the elevations for these angles.

5. Discussion

5.1. Sound Localization First Round: Horizontal Plane

In line with previous studies, the horizontal plane results analysis demonstrated a notable response in frontal and lateral localization accuracy [35]. The presence of visual features in the designed virtual environment, particularly those positioned at the front, may have also contributed to this positive outcome. Although the test was conducted using the same Ambisonics system within the same laboratory as M. Aguilar's [15], due to different methodology strategies in both experiments, as explained below, it was challenging to conduct a direct statistical comparison. However, in frontal positions, the test from this study showed a slightly better performance than that observed in Aguilar's study. Another important result was the non-significant difference at the rear position of 180° according to the V-test. This statistical result implies a pronounced front-back confusion [36]. This perceptual issue was also observed by M. Aguilar, who also found a significant front-back confusion tendency at other sound source positions, both in the front and rear. These results are consistent with the spatial resolution limitations of the algorithm, which are determined by the Ambisonics order and loudspeaker density mentioned in other similar studies [37].

Table 1 illustrates that most responses are concentrated in the expected positions or in their immediate vicinity. This clustering suggests that participants in most of the cases

localized the sound sources accurately, even if not always pointing at the exact expected location. Since angular differences were statistically significant at most positions in the V-test, a consistent performance over the horizontal plane can be assumed. The effective use of 2nd-order Ambisonics confirmed the benefits of HOA, justifying the hybrid solution developed by combining UE and Max, overcoming UE's 1st-order limitations, as will be discussed below.

5.2. Sound Localization Second Round: Horizontal and Vertical Planes

The results of the Hodges–Ajne test revealed that participants could localize sound sources with significant alignment to actual horizontal positions at most spherical angles. This indicates effective horizontal localization, except at 180°, where front-back confusion persisted, suggesting difficulties when trying to locate sounds in the rear position.

The Mardia–Watson–Wheeler test further highlighted the influence of elevation on localization accuracy. Significant differences were noted at 90° and 270°, indicating that participants were sensitive to changes in elevation at both sides, acceptably discerning vertical sound direction. Conversely, the lack of significant differences at 0° and 180° suggests poor auditory discrimination for vertical frontal and rear sources.

Consistent with previous studies on elevation perception [38,39], participants in the second round demonstrated limited ability to discriminate elevated sounds, reflecting the inherent challenges in vertical sound source localization. This difficulty might also be related to the spectral attributes of the sounds used, as sources that are elevated often have a concentrated frequency range in the high-mid spectrum (e.g., birds), which can lead to a bias in perceiving these sounds as coming from above [40].

For the vertical plane assessment, M. Aguilar's data indicated more randomness in response distribution. While both studies only coincided in one position for elevation assessment, direct statistical comparisons were difficult to carry out due to differing methodologies and the specific assessed sound source positions. Nevertheless, the present experiment demonstrated notably accurate lateral elevation localization, which is a significant improvement over previous studies. Additionally, in line with other experiments [18], the order and arrangement of the Ambisonics setup may have influenced the results, suggesting that optimization of these factors could further enhance vertical localization accuracy. These findings further support the enhanced sound localization capabilities afforded by VR and gamification tools.

Additionally, the VR test dynamics were smoother and more interactive, leading to shorter test times and allowing the possibility of considering more audio samples and repetitions.

5.3. User Experience Evaluation

Following the auditory localization tests, participants provided informal verbal feedback, offering valuable insights into their experiences with the VR environment and gamification elements.

The feedback indicated that most participants were highly engaged and enthusiastic, noting that the experiment generated significant interest and involvement. This engagement with the virtual environments contributed to a noticeable reduction in the tests' adjustment period compared to previous listening tests conducted at the LATe lab. The quicker adaptation allowed for a shorter overall test duration, making the process more efficient and less tiring for subjects.

Participants highlighted the rapid learning curve, which reduced training time and improved self-consistency in results. Many participants reported feeling more comfortable during the second part of the test, where the wetland visual environment and Ambisonics field recordings contributed to a more immersive experience. Although this increased comfort did not directly translate into statistical improvements, it is likely that it reduced fatigue and helped participants maintain attention throughout the test.

The VR environment, especially when combined with wetland field recordings, offered participants a sense of landscape familiarity. Interactive features, such as stimulus control and progress tracking, helped keep participants engaged throughout the test.

Moreover, it was clear that the subjective component played a significant role in this experiment. The user experience and interaction with the virtual environment had a noticeable impact, underscoring the complexity of factors influencing auditory perception in VR settings. In future stages, qualitative analyses of the user experience and their interaction with the virtual environment are planned in order to deepen the understanding of these variables and identify possible trends or patterns.

These findings highlighted the positive impact of the VR environment and gamification tools in enhancing participants' engagement and comfort during the tests, suggesting that these features can make localization listening tests more efficient and consistent. However, there is no evidence that these factors had a noticeable impact on results in statistical terms. This suggests that although the VR setup enhanced the user experience, its role in facilitating accurate spatial perception is complex and multifaceted, influenced by factors such as visual dominance, loudspeaker configuration, and users' prior experience with HMDs.

5.4. Hybrid Platform

Another positive outcome of this study was the design and implementation of a hybrid platform that combines UE and Max, expanding the possibilities for higher-order Ambisonics (HOA) reproduction within VR environments. To ensure the stability of the hybrid platform performance, it is crucial that the operating systems of the HMD and the computer used to run the test are compatible and up to date. In this case, addressing this technical issue proved to be a significant challenge in the development and timely implementation of the test.

The decision to use Max and *Spat* spatialization tools proved effective for HOA audio rendering in the tests, offering both stability and detailed control over the spatial attributes of complex Ambisonics encoding and decoding processes. Considering the current limitations of HOA development in UE, these tools hold the promise of being a highly valuable and adaptable addition, paving the way for advanced Ambisonics capabilities in virtual reality and gaming applications.

6. Conclusions

The findings of this study underscore the significant potential of VR tools in advancing the development and implementation of auditory tests. Participants demonstrated improved performance in horizontal sound localization, with results indicating enhanced frontal accuracy and effective lateralization. The setup combining VR design tools with 2nd-order Ambisonics was particularly beneficial, underscoring the importance of high-order Ambisonics reproduction despite the current playback limitations of native UE support.

The use of VR tools in this case contributed to reduced testing time and decreased user fatigue, enhancing usability and consistency in test responses. The rapid adaptation period observed suggests that VR environments can facilitate more efficient testing processes. However, vertical localization results were inconclusive, reflecting the inherent challenges of height perception in human auditory processing, as noted in M. Aguilar's previous study [15].

The results also reveal the potential of VR to create engaging and immersive experiences. Participants reported feeling more comfortable and engaged, particularly in the virtual environments enriched with wetland field recordings. This suggests that VR can help mitigate auditory fatigue and maintain participant attention by offering a sense of familiarity. These insights open new possibilities for implementing innovative spatial audio tools in VR, enabling richer and more interactive experiences.

Future work could explore integrating head-tracking and user movement to further enhance immersion and mobility within a defined listening space. Additionally, further investigation into the impact of Ambisonics field recordings in hybrid audiovisual environments could provide valuable data on the intersection of sound source localization and immersion.

Overall, this research highlights VR's potential as a powerful tool for perceptual and multisensory exploration, extending its applications beyond gaming into significant scientific research and practical uses. Moreover, as it has been discussed in recent works [22,23,41], the use of VR in auditory research has proven to be a field with a rising potential. This challenges the conventional view of VR as merely an entertainment medium, showcasing its broader implications and utility.

Author Contributions: Conceptualization, E.M. and F.O.; methodology, E.M.; statistical analysis, E.M., and R.V.-M.; data curation, E.M. and R.V.-M.; writing—original draft preparation, E.M. and F.O.; writing—review and editing, E.M., F.O., and R.V.-M.; project administration, F.O. All authors have read and agreed to the published version of the manuscript.

Funding: The research that led to this article was funded by the Chilean National Fund for Scientific and Technological Development (FONDECYT) under the project “Herramienta espacio-temporal para aplicaciones creativas y educativas de grabaciones de paisaje sonoro de humedales,” grant no. 1220320. R.V.-M. acknowledges support from ANID FONDECYT de Postdoctorado grant number 3230356.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to the low-risk nature of the research and the use of fully anonymized data. The approving agency for the exemption is Nanjing Forestry University.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data supporting the findings of this study are available within the paper. For additional information, please contact the corresponding author.

Acknowledgments: The authors are grateful to Martín Aguilar and André Rabello Mestre for their support with listening test procedures and manuscript proofreading. Additionally, we extend our thanks to David Valencia, coordinator of the Leufulab laboratory, for his guidance in developing the virtual environments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Barlow, C.; Michailidis, T.; Meadow, G.; Gouch, J.; Rajabally, E. Using immersive audio and vibration to enhance remote diagnosis of mechanical failure in uncrewed vessels. In Proceedings of the Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio 2019, York, UK, 27–29 March 2019.
2. Gorzynski, S.; Kaplanis, N.; Bech, S. A flexible software tool for perceptual evaluation of audio material and VR environments. In Proceedings of the Audio Engineering Society Convention 149 2020, Online, 27–30 October 2020.
3. Hupke, R.; Ordner, J.; Bergner, J.; Nophut, M.; Preihs, S.; Peissig, J. Towards a Virtual Audiovisual Environment for Interactive 3D Audio Productions. In Proceedings of the Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio 2019, York, UK, 27–29 March 2019.
4. Rumsey, F. Perceptual Evaluation: Listening Strategies, Methods, and VR. *J. Audio Eng. Soc.* **2018**, *66*, 301–305.
5. Boukhris, S.; Menelas, B.-A.J. Towards the Use of a Serious Game to Learn to Identify the Location of a 3D Sound in the Virtual Environment. In Proceedings of the Human-Computer Interaction. Interaction Contexts: 19th International Conference 2017, Vancouver, BC, Canada, 9–14 July 2017; pp. 35–44.
6. Wu, Y.-H.; Roginska, A. Analysis and training of human sound localization behavior with VR application. In Proceedings of the Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio 2019, York, UK, 27–29 March 2019.
7. Frank, M.; Perinovic, D. Recreating a multi-loudspeaker experiment in virtual reality. In Proceedings of the Audio Engineering Society Conference: AES 2023 International Conference on Spatial and Immersive Audio 2023, Huddersfield, UK, 23–25 August 2023.
8. Huisman, T.; Ahrens, A.; MacDonald, E. Ambisonics Sound Source Localization with Varying Amount of Visual Information in Virtual Reality. *Front. Virtual Real.* **2021**, *2*, 722321. [[CrossRef](#)]
9. Rees-Jones, J.; Murphy, D. A Comparison Of Player Performance In A Gamified Localisation Task Between Spatial Loudspeaker Systems. In Proceedings of the 20th International Conference on Digital Audio Effects 2017, Edinburgh, Scotland, 5–9 September 2017.

10. Zotter, F.; Frank, M. *Springer Topics in Signal Processing Ambisonics A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*; Springer Nature: New York, NY, USA, 2019.
11. Frank, M.; Zotter, F.; Sontacchi, A. Producing 3D Audio in Ambisonics. In Proceedings of the Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology—Cinema, Television and the Internet 2015, Hollywood, CA, USA, 6–8 March 2015.
12. Valzolgher, C.; Alzhaler, M.; Gessa, E.; Todeschini, M.; Nieto, P.; Verdelet, G.; Salemmme, R.; Gaveau, V.; Marx, M.; Truy, E.; et al. The impact of a visual spatial frame on real sound-source localization in virtual reality. *Curr. Res. Behav. Sci.* **2020**, *1*, 100003. [\[CrossRef\]](#)
13. Native Soundfield Ambisonics Rendering in Unreal Engine | Unreal Engine 5.4 Documentation | Epic Developer Community. Available online: <https://dev.epicgames.com/documentation/en-us/unreal-engine/native-soundfield-ambisonics-rendering-in-unreal-engine> (accessed on 27 July 2024).
14. What is Max? | Cycling '74. Available online: <https://cycling74.com/products/max> (accessed on 27 July 2024).
15. Aguilar, M. Implementación y Evaluación de un Diseño Sonoro en Formato 3D-Ambisonics. Undergraduate Thesis, Universidad Austral de Chile, Valdivia, Chile, 2023.
16. Romigh, G.D.; Brungart, D.S.; Simpson, B.D. Free-Field Localization Performance with a Head-Trackable Virtual Auditory Display. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 943–954. [\[CrossRef\]](#)
17. Moraes, A.N.; Flynn, R.; Hines, A.; Murray, N. Evaluating the User in a Sound Localisation Task in a Virtual Reality Application. In Proceedings of the 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), Athlone, Ireland, 26–28 May 2020; IEEE: New York, NY, USA, 2020; pp. 1–6.
18. Power, P.; Dunn, C.; Davies, B.; Hirst, J. Localisation of Elevated Sources in Higher-Order Ambisonics. *BBC Res. Dev. White Pap. WHP* **2013**, 261.
19. Sochaczewska, K.; Malecki, P.; Piotrowska, M. Evaluation of the Minimum Audible Angle on Horizontal Plane in 3rd order Ambisonic Spherical Playback System. In Proceedings of the 2021 Immersive and 3D Audio: From Architecture to Automotive 2021, Bologna, Italy, 8–10 September 2021.
20. Rudzki, T.; Gomez-Lanzaco, I.; Stubbs, J.; Skoglund, J.; Murphy, D.T.; Kearney, G. Auditory localization in low-bitrate compressed Ambisonic scenes. *Appl. Sci.* **2019**, *9*, 2618. [\[CrossRef\]](#)
21. Thresh, L. A Direct Comparison of Localization Performance When Using First, Third, and Fifth Ambisonics Order for Real Loudspeaker and Virtual Loudspeaker. *J. Audio Eng. Soc.* **2017**, 9864.
22. Piotrowska, M.; Bogus, P.; Pyda, G.; Malecki, P. Virtual Reality as a Tool for Investigating Auditory Perception. In Proceedings of the Audio Engineering Society 156th Convention 2024, Madrid, Spain, 15–17 June 2024.
23. Ramírez, M.; Arend, J.M.; von Gablenz, P.; Liesefeld, H.R.; Pörschmann, C. Toward Sound Localization Testing in Virtual Reality to Aid in the Screening of Auditory Processing Disorders. *Trends Hear.* **2024**, *28*, 23312165241235463. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Otondo, F.; Rabello-Mestre, A. The Soundlapse Project Exploring Spatiotemporal Features of Wetland Soundscapes. *Leonardo* **2022**, *55*, 267–271. [\[CrossRef\]](#)
25. Fernández, J. Diseño y Evaluación Perceptual de un Algoritmo de Espacialización Sonora en Formato Ambisonics. Undergraduate Thesis, Universidad Austral de Chile, Valdivia, Chile, 2023.
26. Carpentier, T. Spat: A comprehensive toolbox for sound spatialization in Max. *Ideas Sonicas* **2021**, *13*, 12–23.
27. LATE | Laboratorio de Arte y Tecnología. Available online: <https://soundlapse.net/latelab/> (accessed on 27 July 2024).
28. Middlebrooks, J.C. Sound localization. In *Handbook of Clinical Neurology*; Elsevier B.V.: Amsterdam, The Netherlands, 2015; Volume 129, pp. 99–116.
29. Batschelet, E. *Circular Statistics in Biology*; Mathematics in Biology; Academic Press: London, UK, 1981; ISBN 0120810506.
30. Mardia, K.V. A Multi-Sample Uniform Scores Test on a Circle and its Parametric Competitor. *J. R. Stat. Soc. Series B Stat. Methodol.* **1972**, *34*, 102–113. [\[CrossRef\]](#)
31. Rayleigh, L., XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1880**, *10*, 73–78. [\[CrossRef\]](#)
32. Hodges, J.L. A Bivariate Sign Test. *Ann. Math. Stat.* **1955**, *26*, 523–527. [\[CrossRef\]](#)
33. Watson, G.S. Circular Statistics in Biology. *Technometrics* **1982**, *24*, 336. [\[CrossRef\]](#)
34. Mardia, K.V. Statistics of Directional Data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1975**, *37*, 349–371. [\[CrossRef\]](#)
35. Hartmann, W.M. Localization and Lateralization of Sound. In *Binaural Hearing: With 93 Illustrations*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 9–45.
36. Wightman, F.L.; Kistler, D.J. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J. Acoust. Soc. Am.* **1999**, *105*, 2841–2853. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Zotter, F.; Pomberger, H.; Noisternig, M. Energy-Preserving Ambisonic Decoding. *Acta Acust. United Acust.* **2012**, *98*, 37–47. [\[CrossRef\]](#)
38. Trapeau, R.; Schönwiesner, M. The Encoding of Sound Source Elevation in the Human Auditory Cortex. *J. Neurosci.* **2018**, *38*, 3252–3264. [\[CrossRef\]](#)
39. Lee, H. Sound Source and Loudspeaker Base Angle Dependency of the Phantom Image Elevation Effect. *J. Audio Eng. Soc.* **2017**, *65*, 733–748. [\[CrossRef\]](#)

40. Lee, H. Psychoacoustics of height perception in 3D audio. In *3D Audio*; Routledge: New York, NY, USA, 2021; pp. 82–98.
41. Stecker, G.C. Using virtual reality to assess auditory performance. *Hear. J.* **2019**, *72*, 20–22. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.