Audio Engineering Society

Convention Paper

Presented at the 144th Convention
2018 May 23 – 26, Milan, Italy

Evaluation of Binaural Renderers: Externalization, Front/Back and Up/Down Confusions

Gregory Reardon[1], Gabriel Zalles[1], Patrick Flanagan[2], and Agnieszka Roginska[1]

[1]*New York University, 35 W. 4th St., New York, NY 10012*
[2]*THX, 1255 Battery St, Suite 100, San Francisco, CA 94111*

Correspondence should be addressed to Patrick Flanagan (`patrick@thx.com`)

ABSTRACT

Binaural renderers can be used to reproduce dynamic spatial audio over headphones and deliver immersive audio content. Six commercially available binaural renderers with different rendering methodologies were evaluated in a multi-phase subjective study. This paper presents the testing methodology, evaluation criteria, and main findings of the first phase. The results of tests on externalization, front/back confusions, and up/down confusions are detailed. Significant effects were found for all three dependent measures.

1 INTRODUCTION

Recent interest in immersive audio has motivated a proliferation of binaural renderers for creating spatial audio content. These technologies render audio scenes for reproduction over headphones. That is, they take a collection of audio waveforms with associated metadata describing the location, reverb characteristics, directivity, etc., of the waveform in virtual space, known together as an *audio object*, and, by leveraging psychophysical features of human hearing, reproduce a 3D sound image over headphones. The location and orientation of audio objects with respect to the user's head can be continuously updated by tracking the user. Audio objects in the virtual scene can then be made to appear as naturally occurring in the user's environment [1]. Within this context there are three main types of sound localization errors that are studied in psychoacoustic literature: *localization*, *externalization*, and *reversal errors*. Of interest in this work are externalization errors and reversal errors.

This work presents results from a portion of a larger subjective experiment on the performance of commercially available binaural renderers. It is beyond the scope of this study to identify the specific renderers that were tested. Rather, the goal here is to design a methodology that can be used for evaluating binaural renderers and to determine the impact these evaluative metrics have in creating an immersive binaural experience.

A three-phase methodology for the evaluation of binaural renderers was presented in a previous work by the authors [2]. A large subjective test was carried out using this methodology. The first phase of the test was concerned with the evaluation of sound localization errors of 3D audio reproduced over headphones. Externalization, front/back and up/down confusions, and localization were assessed individually. This work presents the results of the externalization and front/back and up/down confusions tests. The results from the localization test can be found in [3]. The second phase of the subjective experiment was concerned with evaluating spatial sound quality attributes, such as *naturalness*, *spaciousness*, and *clarity*. The third phase consisted of an overall preference assessment and yielded a forced-choice ranking of renderers. The results of these phases are to follow.

The goals of the subjective experiment at large are twofold. First, the authors seek to understand how the different individual metrics of binaural renderer performance are correlated with overall preference. This will provide an understanding of where to focus improvements in the rendering procedure and how to ameliorate overall renderer performance. Second, the authors want to better understand the variance in performance and take a survey of commercially available binaural renderers. In order to perform both of the above tasks, each test must first be analyzed individually. Further, it also provides an opportunity to evaluate the methodology presented in [2] and indicate any improvements that can be made to the proposed methodology for comprehensive evaluation of binaural renderers.

Related research on the evaluation of perceived externalization and the assessment of reversal errors is first presented. The methodology used for evaluating these sound localization errors in this experiment is reviewed and expounded upon in section 2. The results are presented in section 3. A discussion of the results is found in section 4. Section 5 concludes and outlines future work.

1.1 Externalization

The human auditory system uses two main cues for sound localization based on interaural differences of waves received at each ear. Synthesizing realistic spatial images over headphones was found to not be as simple as imposing these cues onto content. The spatial image tended to degenerate so auditory images appeared internalized, a phenomenon known as *inside-the-head-locatedness* [4]. The externalization of auditory images, such that images appear indistinguishable from spatial images produced by real-world stimuli, has been a large topic of research in psychoacoustics [5]. Initially, it was hypothesized that unsynthesized pinna cues were what caused this spatial image degeneration, but further studies have shown that a number of factors influence externalization.

The first is the use of general versus individualized Head-Related Transfer Functions (HRTFs). HRTFs are composites of interaural differences and pinna cues. They can be gathered from anechoic or semi-anechoic environments by capturing the impulse responses of dummy head microphones (general HRTFs) or microphones placed directly in the ears of a subject (individualized HRTFs). They can also be analytically solved given a 3D model of a head [6]. Personal binaural synthesis has been shown to improve ratings of perceived externalization [7]. HRTFs gathered under non-anechoic conditions are known as binaural room impulse responses (BRIRs). Plausible or realistic reverberation, whether captured in a BRIR or synthesized to complement the HRTFs, has also been shown to improve ratings of externalization [8]. Another factor believed to influence ratings of externalization is the use of head tracking [9, 10]. Though some authors have reported no significant effect of head tracking on externalization [11], the unnatural experience of static binaural content leads to degradation in the spatial image quality. Externalization errors have been found to persist even given dynamic personalized binaural reproduction. Another factor which has recently come under attention to account for this degeneration is the *room divergence effect*. This is defined as a divergence in audiovisual congruence by virtue of differences between synthesized scene and listening environment [12].

1.2 Reversal Errors

A second type of sound localization error particularly endemic in 3D audio reproduced over headphones is that of front/back and up/down reversals. These errors occur along auditory *cones of confusion*. Each auditory cone of confusion is the set of all points on a sphere surrounding an individual's head for which the two main human localization cues, interaural time differences and interaural level differences, are the same. An example of a cone is presented in *Fig. 1*. Within a cone, two types of errors can occur: *front/back reversals* and *up/down reversals*. In such errors, the individual perceives an auditory illusion occurring at a location symmetric to the actual audio event over the frontal plane or transverse plane. Reversal errors over the frontal plane, often referred to as the interaural axis, are known as front-back and back-front confusions. Reversal errors over the transverse plane are known as up-down and down-up confusions [1]. Head tracking is well-known to reduce the prevalence of reversal errors. Even subtle head movements can help to disambiguate sound localization cues [13].

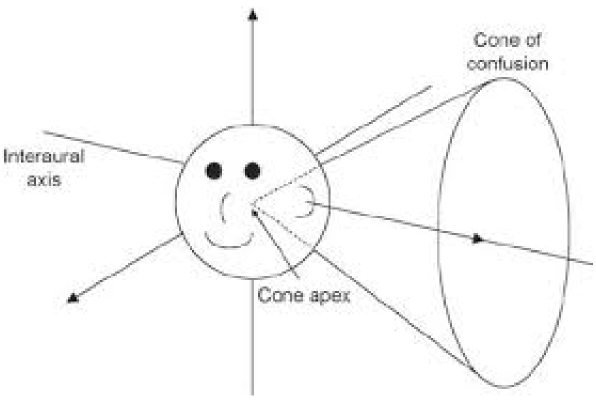

Fig. 1: Graphical depiction of an auditory cone of confusion.

There are two types of front-back reversals based on direction of reversal over the interaural axis. Stimuli presented in front that are reversed to the rear are known as *front-back reversals*, while stimuli presented in the rear reversed to the front are known as *back-front reversals*. Front-back reversals tend to be much more common than back-front reversals [14]. This is typically explained as a part of a primitive survival heuristic that, in the absence of a visual stimuli in front of the listener, defaults to perceiving the location of the acoustic event as occurring behind the individual [1]. With respect to up-down reversals, a stimulus above the head reversed to below the head is known as an *up-down reversal*. A stimulus located below the head reversed to above the head is known as a *down-up reversal*. Previous studies have found no particular statistical bias towards either reversal direction over the transverse plane [15].

2 METHODOLOGY

2.1 Rendering Procedure, Stimuli, and Presentation

Six different commercial renderers were tested comparatively. The renderers are labelled from 00 - 05. Renderers 00, 01, and 05 render binaural audio using higher-order Ambisonics (HOA). Renderers 03 and 04 use first-order Ambisonics (FOA). Renderer 02 uses direct virtualization through HRTFs. Three different stimuli were used to assess externalization, front-back and up-down confusions for these different binaural renderers. The stimuli were two-second mono drum loops of different styles created in Pro Tools. These stimuli were output at 48 kHz sampling rate and 24 bit depth. A decision was made against testing those stimuli traditionally found in psychoacoustic literature, such as noise or infrapitch sound. This selection reflects a desire to understand how the renderers perform in commercial settings and also provides relatively broadband signals that are able to exploit the full range of auditory cues.

For each renderer, the stimuli were processed in the audio scene at a chosen distance of one meter from the listener at various azimuths and elevation angles. Though each of renderers supported headtracking in their native application, for the purposes of the experiment, the content was head-locked. Thus a discrete set of over one thousand stimuli was rendered as static binaural audio to be used in the first phase of the experiment. Each subject was presented a subset of these stimuli. In an attempt to evaluate the base rendering engine of each binaural renderer, all room information was turned off. This included both room reverb and early reflections, for those renderers that supported such a property. This also permitted more uniform comparison between renderers. All other renderer properties were set to their highest quality.

A total of sixty-eight subjects participated in the test. Stimuli were presented over circumaural stereophonic headphones (Sennheiser HD-650) in a soundproof booth (NYU Dolan Isolation Booth). Custom software was used to administer the experiment and collect data. Subjects indicated their responses directly on a graphical user interface and had the option to replay any stimuli before submitting a response and moving on to the next trial. For each trial, subjects were given a comment box to provide feedback on that specific trial.

2.2 Externalization Test

In the externalization portion of the test, each subject performed twenty-four trials - 6 training and 18 testing. In the test trials, each renderer and each stimulus type was shown once. Defined here and for the rest of the paper, 0° azimuth refers to directly in front of the subject and with azimuth increasing in a clockwise direction (90° refers to the right of the subject and 270° refer to the left of subject). In this test all stimuli presented were located on the horizontal plane (0° elevation). The

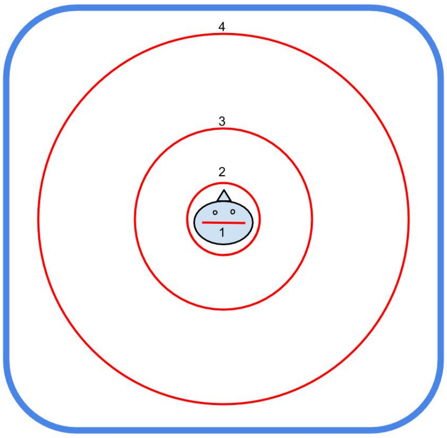

Fig. 2: Graphical representation of the discrete levels of externalization tested.

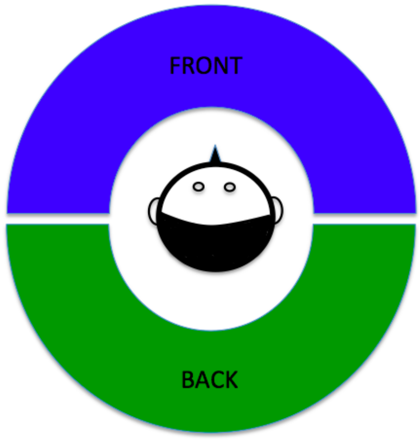

Fig. 3: Accompanying graphic used to assess front-back confusions.

set of azimuths in this test were all locations at 10° azimuth increments (0°, 10°, ...340°, 350°). In each trial, four stimuli were presented. The first presentation was always a reference unprocessed stimuli. The following three were spatialized versions of that reference stimuli, drawn at random from the set of all possible azimuths. Subjects were asked to rate on a scale from one to four, where one was "inside-the-head," and four was "far away from the head, external in space", how far away the subset of spatialized stimuli appeared as a whole. A graphic accompanied this verbal description and is pictured in *Fig. 2*. The goal of presenting multiple spatialized stimuli at different azimuths within each trial is to improve the consistency of externalization ratings. Given that ratings of externalization will be dependent on the location of the stimuli with respect to the head [14], presenting individual stimulus would increase the variance of results.

2.3 Front/Back and Up/Down Confusion Tests

In the front/back test, subjects performed eighteen trials. Each renderer and each stimulus was shown once per subject. In each trial, a cone was first selected. On the horizontal plane, each cone of confusion is associated with two azimuth locations. The associated azimuth pairs on the horizontal plane that were tested are as follows:

$$Pairs = [T, 180^\circ - T] \tag{1}$$

where $-60^\circ <= T <= -20^\circ$ and $20^\circ <= T <= 60^\circ$ and T is a multiple of 10. A pair of spatialized stimuli were selected from the above set, the order randomized, and the stimuli presented. Subjects were asked to indicate the location of the second sound in the pair as either in front of or behind their head. Subjects are thus asked to make relative, as opposed to absolute judgments, of sound source location. The motivation for which is to improve the consistency of judgements. The graphic that accompanied this description is included in *Fig. 3*.

The up/down test was similar. Subjects performed eighteen trials, one for each renderer and each stimulus. In this case, reversals over the transverse plane are of interest. The decision was made to test elevations of +30° and -30°. In each trial, an azimuth was selected at random from the following set: (20°, 30°, ... , 60°, 120°, 130°, ..., 160°, 200°, 210°, ... , 240°, 300°, 310°, ... , 340°). The pair of spatialized stimuli located at that azimuth with elevation +30° and -30° were grabbed from the repository of stimuli, the order randomized, and the stimuli presented. Subjects were asked to indicate the location of the second sound as either above the head or below the head. A similar, but appropriately altered, graphic as that presented in *Fig. 3* accompanied this description.

3 RESULTS

A 6 x 3 univariate repeated-measures analysis of variance (ANOVA) was conducted for each of the dependent measures: externalization, front/back reversals,

	00	01	02	03	04	05
00	-----					
01	0.752	-----				
02	0.457	0.002*	-----			
03	< 0.001*	< 0.001*	< 0.001*	-----		
04	1.000	0.381	1.000	< 0.001*	-----	
05	0.024*	< 0.001*	1.000	< 0.001*	0.247	-----

Table 1: Externalization - significance levels for multiple comparisons of renderers

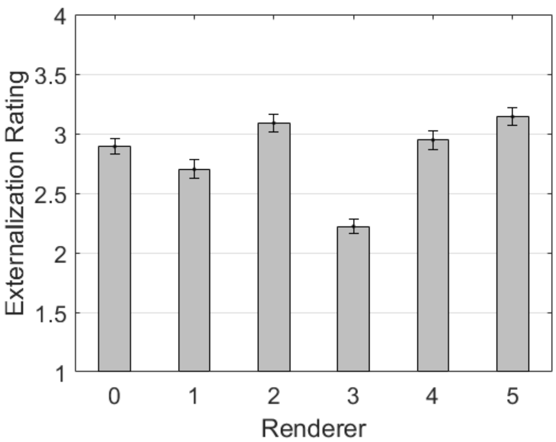

Fig. 4: Externalization - main effect for renderer (mean and standard error bars)

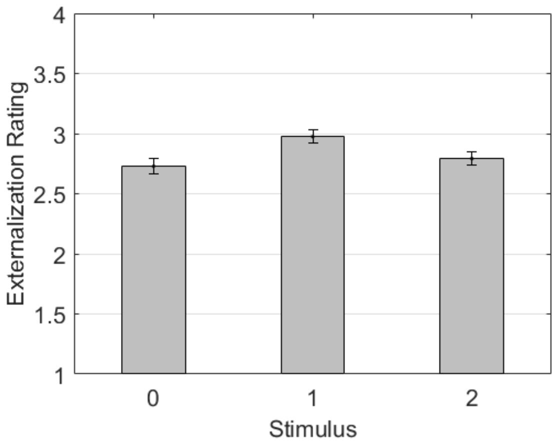

Fig. 5: Externalization - main effect for stimulus (mean and standard error bars)

up/down reversals. Renderers were labeled 00 - 05; and stimuli were labeled 0 - 2. A further breakdown of reversals by virtue of type (front-back versus back-front and up-down versus down-up) is also presented. A significance level of 0.05 and was used for all statistical tests. All subjects included in the results self-reported as having normal hearing.

3.1 Externalization

As discussed, ratings of externalization were performed on a discrete 1-4 scale, with 4 being "far away from the head, external in space", and 1 being "inside the head." Sixty-eight subjects participated in the externalization test. The 6 training trials were excluded from the dataset. The main effect for renderer was found to be significant, F(5,3) = 28.414, $p < 0.001$. Sphericity assumptions were almost violated ($p = 0.069$), but the main effect was still significant even with the Geisser-Greenhouse correction ($p < 0.001$). These results are pictured in *Fig. 4*. A multiple comparisons test with a Bonferroni confidence level correction was performed to better understand how each renderer performed against each of the others. The significance levels of this comparison are pictured in *Table 1*. Renderer 03 was found to be significantly different than each of the other renderers. Significance between other pairs varies. It is not the case that each renderer is statistically different from the others. Some perform very similarly.

The main effect for stimulus was also found to be sig-

	00	01	02	03	04	05
00	-----					
01	0.437	-----				
02	1.000	0.826	-----			
03	1.000	0.035*	1.000	-----		
04	< 0.001*	< 0.001*	< 0.001*	< 0.001*	-----	
05	1.000	1.000	1.000	1.000	< 0.001*	-----

Table 2: Front/Back Confusions - significance levels for multiple comparisons of renderers

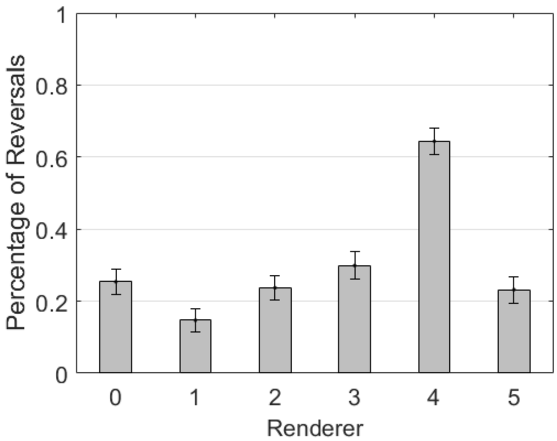

Fig. 6: Front/Back Confusions - main effect for renderer (mean and standard error bars)

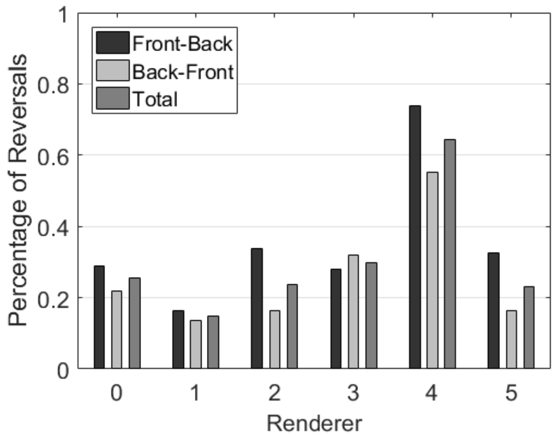

Fig. 7: Front/Back Confusions - reversal error type (percentage of total number of trials possible to reverse in each direction)

nificant, F(3,6) = 6.777, $p = 0.02$. The ANOVA also indicated a significant interaction between renderer and stimulus, F(1, 10) = 10.00, $p = 0.001$. Mean and standard errors in externalization for each stimulus are presented in *Fig. 5*. A significant omnibus test does not give an understanding of why this is the case; multiple comparisons were also analyzed in this circumstance. The post hoc test revealed that stimulus 1 returned significantly higher rating of perceived externalization when compared to either stimulus 0 ($p = 0.004$) or stimulus 2 ($p = 0.018$). There was no significant difference between stimulus 0 and 2.

3.2 Front/Back Confusions

Front-back confusions were assessed by having subjects indicate whether the second sound in the pair of stimuli was located as either "behind the head or "in front of the head." An initial ANOVA test was first run using a subset of the data (twenty-nine subjects). The ANOVA indicated a significant effect for renderer, F(5, 3) = 17.536, $p < 0.001$, but not for stimulus. No significant interaction between renderer and stimulus was found.

Given that stimulus was not found to be significant, the whole dataset of fifty-nine subjects who participated in the test was used to establish the overall mean reversal rates for each renderer. In this ANOVA, the main effect for renderer was once again significant, F(5,3) = 24.956, $p < 0.001$. These results are found in *Fig. 6*. A post hoc multiple comparisons test between renderers with a Bonferroni correction was once again performed.

	00	01	02	03	04	05
00	-----					
01	0.003*	-----				
02	0.052	1.000	-----			
03	0.001*	< 0.001*	< 0.001*	-----		
04	1.000	< 0.001*	0.003*	0.133	-----	
05	0.447	< 0.001*	< 0.001*	1.000	1.000	-----

Table 3: Up/Down Confusions - significance levels for multiple comparisons of renderers

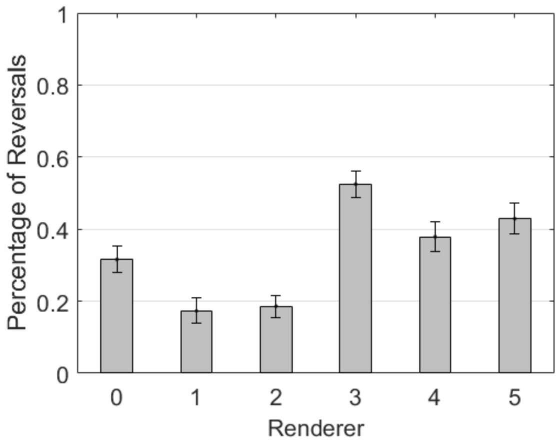

Fig. 8: Up/Down Confusions - main effect for renderer (mean and standard error bars)

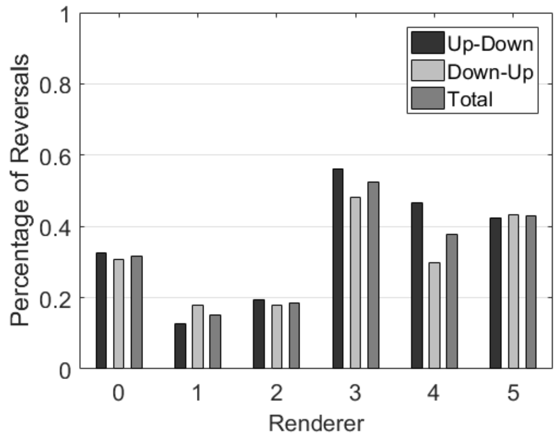

Fig. 9: Front/Back Confusions - reversal error type (percentage of total number of trials possible to reverse in each direction)

These results are presented in *Table 2*. Renderer 04 performed extremely poorly in this test. With a mean reversal rate greater than fifty percent.

Fig. 7 presents a breakdown of the type of reversal errors for each renderer. The graph is categorized by frontal stimuli reversed to the rear (front-back reversals), rear stimuli reversed to the front (back-front reversals), and total reversal errors as a percentage of total trials possible to reverse in either direction. Front-back reversals were much more common. Renderer 04 once again stands out in this respect, with a consistent bias for reversing frontal stimuli to the rear.

3.3 Up/Down Confusions

Similarly to the front-back test, subjects assessed whether the location of the second sound in a pair of stimuli was "above the head" or "below the head." Fifty-nine subjects participated in the up-down confusion test. The ANOVA indicated that there was a significant main effect for renderer, F(5,3) = 17.333, $p < 0.001$. Significant effects were not found for stimulus or for renderer-stimulus interactions. The mean reversal rates for each renderer are presented in *Fig. 8*. The results of the post hoc multiple comparisons test between renderers with a Bonferroni correction is presented in *Table 3*. Renderer 04 did not perform nearly as poorly as in the front-back confusion test, but it did cluster with renderers 03 and 05 as the weakest performers in the up/down confusions test.

Fig. 9 breaks down the performance of the renderers in the up/down test by type of reversal error. The graph displays up-down reversal errors, down-up reversal errors, and total reversal errors as a percentage of the total number of trials to reverse in either direction. There does not appear to be consistent bias in either direction of reversal.

4 DISCUSSION

It is possible to evaluate these results in isolation from the rest of the subjective experiment in a few ways. The most important result is that the performance of commercial renderers in terms of externalization, front/back and up/down confusions is, statistically speaking, different. For each of the individual tests, the main effect for renderer was significant. In each test, different renderers excelled, indicating that specific rendering techniques result in trade-offs in performance. There does not appear to be defined clustering of results when grouped by spatialization method. But the FOA renderers, renderer 03 and 04, do, generally, perform poorly when compared to the other renderers. Specifically, the poor performance of renderer 03 in the externalization test (*Fig. 4*) and of renderer 04 in the front-back test (. *6*) stand out.

An interesting result in the externalization test was that the ANOVA indicated a significant main effect for stimulus. This is consistent with comments received during testing. Subjects noted that the unprocessed stimulus 1 appeared more externalized than the the other two unprocessed stimuli. This stimulus had more broadband frequency content than the other stimuli. The dependency of perceived externalization on stimulus (broadband, speech, etc.) is well-documented [11] and the differences in frequency content of the stimuli likely led to this result. In comments, subjects also noted the variance in perceived externalization of spatialized stimuli within the same trial. For instance, in a single trial a subject might have been presented a set of stimuli located at 0°, 180°, and 270° azimuth. The stimuli at 0° and 180° would appear close to the head, while the stimulus at 270° would appear much further, externalized in space. More generally, externalization is azimuth dependent and perceived externalization tends to be worse for stimuli presented at the front or back of the head when compared with stimuli presented at the sides of the head. This has been reported in previous studies [12, 14]. For subjects in this experiment, within a trial, it was occasionally difficult to come a single judgement for perceived level of externalization for this reason. The authors attempted to control for this across all subjects by presenting three stimuli from random azimuths, but, as seen in *Table 1*, significant differences between most renderers in the multiple comparisons test were not found. The lack of significance between renderers and the comments received during testing highlight that there might be improved methods for testing externalization. Alternatively, the authors considered using more than three spatialized stimuli per trial, played in a trajectory around the head.

As seen in *Fig. 7* there does appear to be consistent bias for front-back reversals for each of the renderers. This has been reported by other authors [1]. Renderer 03 was the only renderers in which back-front reversals were more prevalent. Presenting either front-back to back-front trajectories with repetition was meant to improve the accuracy of localization judgements. Even with this methodology, renderer 04 had greater than fifty percent mean reversal rate. The other renderers performed similarly and significant differences in the multiple comparisons test, barring renderer 04, were found only between renderer 01 and 03.

Consistent with the findings in [15], the up/down confusions test did not reveal any particular bias for type of reversal error, either up-down or down-up. With respect to mean reversal rates, only two renderers had below twenty percent reversal rates. Renderers 01 and 02 stood out in this test and, looking more generally across all three tests, perform strongly on all metrics examined in this work. These renderers are a HOA renderer and a direct virtualization renderer, respectively.

5 CONCLUSIONS AND FUTURE WORK

A large multi-part subjective experiment was conducted. sixty-eight paid subjects participated. The results from the externalization, front-back and up-down confusions tests of the experiment are presented. Significant effects were found in all three tests. Renderer 01 and 02 generally perform strongly, while the performance of renderers 03 and 04 (the FOA renderers) are generally poor. No renderer performed best in all three tests indicating that there are tradeoffs in baseline performance due to rendering methodology.

The goal of the larger subjective experiment is to understand how these various judgements, such as externalization, front/back and up/down confusions, along with the metrics used in the other phases of the test, are related to assessments of preference for commercially available binaural renderers. That being said, it is possible to still gain meaningful information from the individual tests. The multiple comparisons tests give an understanding of how the methodology performs when discriminating between pairs of renderers in the experiment. This is important for generalizing the methodology in experiments evaluating less than six binaural renderers. Such, or a similar, context is of interest to the authors during future work. The results also indicate that there might be possible improvement to the presentation of stimuli and the selection of stimuli in the externalization test. Initial results are encouraging in the context of the larger subjective experiment which aims to better understand spatial audio perception in binaural renderers.

6 ACKNOWLEDGEMENTS

The authors would like to thank THX Ltd for their support on this research. Special thanks to Dr. Johanna Devaney's statistics guidance.

References

[1] Begault, D. R. and Trejo, L. J., "3-D sound for virtual reality and multimedia," 2000.

[2] Reardon, G., Calle, J. S., Genovese, A., Zalles, G., Olko, M., Jerez, C., Flanagan, P., and Roginska, A., "Evaluation of Binaural Renderers: A Methodology," in *Audio Engineering Society Convention 143*, Audio Engineering Society, 2017.

[3] Reardon, G., Genovese, A., Zalles, G., Flanagan, P., and Roginska, A., "Evaluation of Binaural Renderers: Localization," in *Audio Engineering Society Convention 144*, Audio Engineering Society, 2018.

[4] Mills, A. W., "Lateralization of High-Frequency Tones," *The Journal of the Acoustical Society of America*, 32(1), pp. 132–134, 1960.

[5] Hartmann, W. M. and Wittenberg, A., "On the externalization of sound images," *The Journal of the Acoustical Society of America*, 99(6), pp. 3678–3688, 1996.

[6] Gumerov, N. A., O'Donovan, A. E., Duraiswami, R., and Zotkin, D. N., "Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation," *The Journal of the Acoustical Society of America*, 127(1), pp. 370–386, 2010.

[7] Werner, S., Klein, F., Mayenfels, T., and Brandenburg, K., "A summary on acoustic room divergence and its effect on externalization of auditory events," in *Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference on*, pp. 1–6, IEEE, 2016.

[8] Durlach, N. I., Rigopulos, A., Pang, X., Woods, W., Kulkarni, A., Colburn, H., and Wenzel, E., "On the externalization of auditory images," *Presence: Teleoperators & Virtual Environments*, 1(2), pp. 251–257, 1992.

[9] Brimijoin, W. O., Boyd, A. W., and Akeroyd, M. A., "The contribution of head movement to the externalization and internalization of sounds," *PloS one*, 8(12), p. e83068, 2013.

[10] Werner, S., Götz, G., and Klein, F., "Influence of Head Tracking on the Externalization of Auditory Events at Divergence between Synthesized and Listening Room Using a Binaural Headphone System," in *Audio Engineering Society Convention 142*, Audio Engineering Society, 2017.

[11] Begault, D. R., Wenzel, E. M., and Anderson, M. R., "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *Journal of the Audio Engineering Society*, 49(10), pp. 904–916, 2001.

[12] Werner, S. and Klein, F., "Influence of Context Dependent Quality Parameters on the Perception of Externalization and Direction of an Auditory Event," in *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*, Audio Engineering Society, 2014.

[13] Thurlow, W. R. and Runge, P. S., "Effect of induced head movements on localization of direction of sounds," *The Journal of the Acoustical Society of America*, 42(2), pp. 480–488, 1967.

[14] Begault, D. R. and Wenzel, E. M., "Headphone localization of speech," *Human Factors*, 35(2), pp. 361–376, 1993.

[15] Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L., "Localization using nonindividualized head-related transfer functions," *The Journal of the Acoustical Society of America*, 94(1), pp. 111–123, 1993.