DOI: 10.61782/fa.2023.0633

forum acusticum 2023

EVALUATING THE PLAUSIBILITY OF NON-INDIVIDUAL HEAD-RELATED TRANSFER FUNCTIONS IN ANECHOIC CONDITIONS

Tim Lübeck[1,2*] **Christoph Pörschmann**[1]
[1] Institute of Computer and Communication Technology, TH Köln, Germany
[2] Audio Communication Group, TU Berlin, Germany

ABSTRACT

Plausibility has become a popular paradigm for perceptually evaluating the binaural reproduction of spatial sound scenes for virtual and augmented reality applications. It refers to a listener's belief in the realism of a virtual sound source where no direct real-world reference is available. Several studies have investigated the plausibility of various virtual acoustic environments. However, no study examined the plausibility of single non-individual head-related transfer functions (HRTFs). HRTFs form the basis of most currently used binaural reproduction methods, either based on simulation, measurements, or parametric sound field descriptions. Therefore, it is an essential research question whether non-individual HRTFs can produce a plausible virtualization of sound sources.

In this paper, we present the results of a listening experiment evaluating the plausibility of HRTFs measured with a Neumann-KU100 artificial head. We conducted an experiment in an anechoic chamber in which participants were asked whether a sound originated from one of four loudspeakers or headphones. The loudspeakers were arranged in a circle with a radius of 1.62 m around the listener and covered with acoustically transparent curtain. The virtual sources were created by convolution with HRTFs and corresponding loudspeaker transfer functions. The results indicate that the participants could not reliably discriminate between real and virtual reproduction.

Keywords: *plausibility, head-related transfer functions, parametric spatial audio*

**Corresponding author*: *tim.luebeck@th-koelen.de.*

1. INTRODUCTION

In Extended Reality (XR) applications, it becomes increasingly relevant to present virtual sound sources together with the real acoustic environment. To perceptually evaluate the rendering of the virtual acoustic environment (VAE) for XR, the paradigm of plausibility has become very popular. While the more traditional paradigm of Authenticity [1, 2] refers to the perceived indistinguishability of the virtual source and a direct *explicit* reference, plausibility refers to the listener's *inner* reference and personal experience. Slater [3] firstly introduced the term plausible illusion and discussed the importance of plausibility for the perception of virtual realities. Lindau et al. [4] defined the plausibility as the agreement with the listener's expectation towards a corresponding real acoustic event, which matches the definition of external plausibility by Hofer et al. [5]. This means that a virtual sound source can be considered plausible if the listener believes that it could be a real source in the real environment. This makes the plausibility an appropriate and popular paradigm for the perceptual evaluation of VAEs for XR applications. Lindau et al [4] further introduced a method based on the signal detection theory (SDT) to assess the plausibility of virtual acoustics, which has become an established method in recent research. Listeners are randomly presented with either a real or a virtual sound source. In a yes/no task, they have to decide which version they were listening to.

A common method for the virtualization of sound sources in acoustic spaces is headphone-based dynamic binaural rendering, which is usually based on head-related transfer functions (HRTFs). Auralization, either based on simulation [6], synthesized from microphone array captures [7, 8], or rendered with parametric approaches [9, 10], usually integrates HRTFs into the processing chain. For consumer applications, mostly non-individual HRTFs are available for the synthesis of the VAE. In the context of developing immersive XR environments, it is thus an essential research question whether non-individual HRTFs can produce plausible virtualization of sound sources.

A number of studies already evaluated the plausibility of VAEs. Most of them investigated the auralization of reverberant environments using different rendering methods, such as synthesis based on BRIR measuremens [4, 11], synthesis based on room acoustic simulation [12], synthesis based on microphone array captures [13], or synthesis based on parametric approaches [14, 15]. Some of them were conducted in a VAE with 6 degrees-of-freedom (6 DoF) [11], some of them with 3 DoF. The specific rendering details, and different room acoustics may all influence the perceived plausibility in different ways and it is not clear which details affect the plausibility to what extend. Those details could be technical aspects such as tracking latency, artifacts due to interpolation between measurements, or spatial aliasing in synthesis from microphone arrays. Furthermore, perceptual aspects like room divergence effects [16] may also affect the perceived plausibility. To the best of our knowledge, only two studies have evaluated the plausibilty in anechoic conditions, and hence, excluded any influence of the specific algorithm for the synthesis of room acoustics. Arend et al. [17] presented a listening experiment to investigate the influence of near-field cues on plausibilty. Here the reference presentation was not a real loudspeaker, but a headphone-based presentation consisting of HRTFs synthesized with distance variation functions. Thus, Arend et al. did not strictly apply the method proposed by Lindau et al. [4]. It was shown that there is no evidence that near-field cues enhance the plausibility of non-individual binaural rendering. The experiment presented by Oberem et al. [18] compared two different individual binaural reproduction methods in which HRTFs were measured in the participants' ear canal. Both individual renderings were considered plausible augmentations of virtual sound sources.

In the present study, we assessed the perceived plausibility of non-individual HRTFs. It is part of a series of listening experiments, in which we plan to successively investigate which aspects of a VAE influence its plausibility. We conducted a listening experiment in anechoic conditions according to the test design proposed by [4]. We presented virtual and real sound sources from the left, front, right, and back. To focus only on the auditory cues, all loudspeakers were covered with an acoustically transparent curtain. The virtual sound sources were synthesized by convolution with generic HRTFs measured on an Neumann KU100 dummy head, and the frequency response of the loudspeaker. Although, the participants remained seated during the experiment, the VAE were rendered with 6 DoF. Hence, not only the listener's head orientation, but also left/right or up/down movements were considered in the binaural synthesis.

2. PARAMETRIC RENDERER

The (position) dynamic binaural rendering was performed with a renderer implemented in Max/MSP. The listener's head orientation and position as well as the position of the loudspeakers are tracked with an Optitrack system. According to the relative direction of the listener's head and the position of the loudspeakers, the corresponding HRTFs are convolved with the dry audio content and adjusted in level according to the $\frac{1}{r}$ distance law. Additionally, the convolved HRTF is filtered with the loudspeaker frequency response for the respective angle. The renderer is generally capable to presenting a 6 DoF dynamic binaural synthesis of reverberant rooms, including synthesis of early reflections according to an image source model, and reverberation. In this study, however, it only rendered the direct sound.

3. METHODS

3.1 Setup and Materials

We used the renderer described in Section 2 in Max/MSP Version 8.1.5. The listeners' head orientation and position was tracked with six cameras from an Optitrack Flex 13 system with the Motive Software Version 2.2.0. The tracking data were then send via OSC through a python middleware to the Max/MSP software. As a digital analog converter we used an RME UFX 2 interface with a sampling rate of 48000 Hz and a buffer size of 256 samples. As headphones we used the extraaural AKG K1000 headphones in the open configuration to reduce the shadowing effect of the headphones during the playback of the real sound sources. Accordingly, the employed generic HRTF set were measured on a Neumann KU100 dummy head wearing AKG K1000 headphones [19]. The HRTF set is available on a 2702 sampling point Lebedev grid in SOFA format [1]. The binaural chain was equalized with headphone compensation filters from [20]. The real sound sources were presented with Genelec 1029a loudspeakers. The corresponding loudspeaker directivity was measured on a N=44 Lebedev grid in the anechoic chamber at TH Köln. It was then resampled in the spherical harmonics domain to obtain a horizontal set of 360 directions in 1° steps. Hence, during the reproduction the loudspeaker directivity was only updated with respect to the relative horizontal angle, which is a fair assumption since the participants sat on a chair during the experiment with their head approximately on the same height as the acoustical center of the loudspeakers. The directivity was further octave band smoothed exported as 2048 tabs minimum phase FIR filters. All loudspeakers were placed at a distance of 1.62 m of the listener at azimuth directions ϕ = 90° [2], ϕ = 0°, ϕ = 270°, and ϕ = 315°, and arranged such that they all faced the listener. This allows to evaluate lateral and frontal sound incidences, as well as backward incidences with slight binaural cues. Although the participants remained seated during the experiment they were allowed to make translational head movements. The loudspeaker directivity and direct sound level were updated accordingly. Due to the loudspeaker distance of 1.62 m we could neglect any near-field effects such as ILD modifications [21]. The test was conducted in the anechoic chamber at the TH Köln. As anechoic dry audio we used 25 different samples including drums, piano, guitar, saxophone, cello, female, and male speech. Since the anechoic chamber at the TH Köln can no longer be considered anechoic below 200 Hz, we high-pass filtered all test signals with a cut-off at 200 Hz, to ensure equal conditions for virtual and real presentation.

3.2 Procedure

The experiment was performed according to the procedure proposed by Lindau et al. [4]. The participants sat on a swivel chair in the anechoic chamber at TH Köln, and were provided with a MIDI controller. To ensure that all participants had a similar expectation of the real acoustic event inside the anechoic chamber, a guitar sample was played through the frontal real loudspeaker as a training stimulus at the beginning of the experiment. Thereby, the participants were aware that this was a playback through the real loudspeaker. This test sample was not one of the samples played back during the actual experiment. Upon entering the room, participants could see all loudspeakers. After the training stimulus, the experimenter turned off the light such that no speaker was visible. Then, 100 trials were played back either through the real loudspeakers, or virtually through the headphones. The number of virtual and real sources was slightly unbalanced, i.e. 47 and 53, 48 and 52, or 49 and 51, to avoid participants making assumptions about the number of correct and false ratings. Furthermore, it was ensured that a real source was never followed by a virtual one at the same position, or vice versa. After each trial, the participants had to answer the question: "Did the sound come from one of the loudspeakers" by pressing the corresponding button on the MIDI board. It took the subjects between 10 to 15 minutes to finish all trials.

3.3 Participants

11 subjects (mean age: 36.18) took part in the experiment. One of them was female, and 10 of them were male. 5 of them had experience in binaural technology and can be considered experienced listeners.

3.4 Data Analysis

We evaluated the results with the SDT. For a detailed derivation the reader is referred to Stanislaw and Todorov [22] or Lindau et al. [4]. There are four possible outcomes of the yes/no paradigm: hit (virtual reproduction is correctly indicated), false alarm (a real source was indicated as virtual), miss (a virtual source was indicated as real), or correct rejection (a real source was indicated as real). The rate of hits (p_{hit}) and false alarms (p_{fa}) fully describe the performance of the participant. The sensitivity d is a measure of how well partcipants could discriminate between real and virtual reproduction and can be estimated from the hit and false alarms rates with $d' = Z(p_{hit}) - Z(p_{fa})$, with $Z(p)$ the inverse cumulative normal distribution. An estimated sensitivity of $d' = 0$ indicates no difference between virtual and real reproduction, and hence perfect plausibility. $d' > 0$ indicates that participants were able to determine if the presentation was real or virtual. It is often easier to interpret the difference in terms of forced choice values, which is why d' can be converted to p_{2AFC} values with $p_{2AFC} = \Phi(d'/\sqrt(2))$ with the cumulative standard normal distribution $\Phi(.)$. Besides the estimated sensitivity d', the bias of the participants is an important parameter in the SDT. It indicates if the participant has a tendency to believe in the real-

[1] `https://doi.org/10.5281/zenodo.3928465`

[2] ϕ denotes the horizontal angle ranging from 0° to 359°, and θ the colatitude angle, ranging from 0 to 180.

ness of the sound source, or to have the tendency towards virtual. There are multiple measures of the bias. In this work we decided to use β, which can be calculated with $\beta' = e^{(-Z(p_{hit})*\frac{Z(p_{hit})}{2}+Z(p_{fa})*\frac{Z(p_{fa})}{2}}$. $\beta > 1.0$ indicates a bias towards real, $\beta < 1.0$ a bias towards virtual. A point of discussion in the evaluation, is the critical sensitivity d_{min} at which an auralization can be considered plausible. Perfect plausibility would mean achieving a sensitivity of $d' = 0$. Lindau et al [4] suggested a $d_{min} = 0.1777$, which corresponds to $p_{2AFC} = 55\%$. They hence allow to exceed the pure guessing rate of 50% by a maximum of 5%. Other studies, use a less strict criterion and use a $p_{2AFC} = 75\%$, which is the point of subjective equality on the psychometrical function [12]. For performing SDT analysis, we used the d' method from the R psycho package (version 0.6.1).

4. RESULTS

A first overview of the results is shown in Fig. 1. It depicts the individual estimated sensitivities d', the corresponding mean values across subjects d'_{avg}, the mean bias β'_{avg}, and the 95% between-subject confidence intervals (CIs). In the first part of the analysis we refer to a critical d_{min} = 0.177, which is indicated as a dashed line. It can bee seen, that the average sensitivity across subjects d'_{avg}, is slightly above d_{min}. However, the CIs overlap, which suggests that d'_{avg} does not significantly differ from d_{min}. The bias β' shows that participants had the tendency towards the answer real.

A Hochberg corrected Shapiro-Wilk test showed no violation of the assumption of normality distribution which is why we used t-tests for the following analysis. An one-sample t-test against $d_{min} = 0.177$ yielded no significant difference of d'_{avg} and d_{min} ($t(10) = 1.8$, $p <$ .103). However, in frequentist statistics, non-significant results are no proof of the null-hypothesis. Therefore, we calculated the Bayes factor (BF_{01}) for the one-sample t-test [23]. We computed a factor of BF_{01} = 0.1544 which indicates that its more likely that participants are within the range of 55% guessing rate, than above.

Moreover, we analyzed the results with respect to speaker position. The corresponding individual d', the average across subject d'_{avg}, β'_{avg}, and the corresponding between-subject CIs are depicted in Fig 2. It can be observed, that for all speakers, the CIs again overlap with the $d_{min} = 0.177$. In the second part of the analysis we refer to more strict critical distance of $d_{min} = 0$, to investigate differences between the rating of each loudspeaker. We again performed t-tests against d_{min} and calculated the Bayes factors. They revealed that for the lateral speakers at the left and right position, we could not reveal a difference from 0. For the left speaker ($t(10) = 1.565$, $p = 0.15$), $BF_{01} = -0.103$), right speaker ($t(10) = 1.768$, $p = 0.108$, $BF_{01} = 0.122$). For the frontal and back speakers the t-test revealed significant differences from 0. Frontal speaker ($t(10) = 2.384$, $p = 0.038$, $BF_{01} = 0.93$), and lateral speaker ($t(10) = 2.734$, $p = 0.021$, $BF_{01} = 1.464$). The bias' β' for the left and right speaker are notably above 1.0, for the frontal and back speaker, the CIs of the bias slightly cross the 1 limit.

Table 1. Mean values of the estimated d' across subjects and corresponding 95% between-subject confidence intervals for all speakers, and each speaker separately.

Speaker	d'_{avg}
All	0.34 ± 0.277
left	0.21 ± 0.3
front	0.46 ± 0.43
right	0.30 ± 0.386
back	0.34 ± 0.28

5. DISCUSSION

The results of the presented listening experiment suggest that non-individual HRTFs can produce plausible virtualization of sound sources. All participants reported that they could not reliably discriminate between headphone and loudspeaker presentation and could only make a guess. This is supported by the results and statistical analysis. Allowing a tolerance from pure guessing rate of 5%, the auralization can be considered plausible. For lateral sound sources, we could not even find a difference to the pure guessing rate of 50%. Some studies have shown that non-individual HRTFs may lead to higher front/back confusion rates [24, 25], which may explain the smaller d'_{avg} of the lateral compared to the front and back sources. Another explanation could be the influence of the AKG K1000 in open configuration. The rear speaker, which radiates directly onto the membrane of the headphones in the open configuration, can cause comb filtering effects

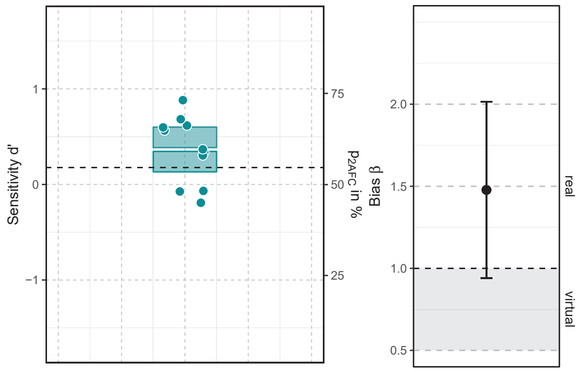

Figure 1. Left: Individual sensitivities d' indicated as dots, the mean across subjects d'_{avg} indicated as the white line, and the 95% confidence intervals indicated as the top and bottom of the box. Right: Mean bias β'_{avg} indicated as a black dot, and 95% confidence intervals indicated as error bars. The dashed line shows the critical d_{min} = 0.177, which corresponds to a P_{AFC} = 0.55%.

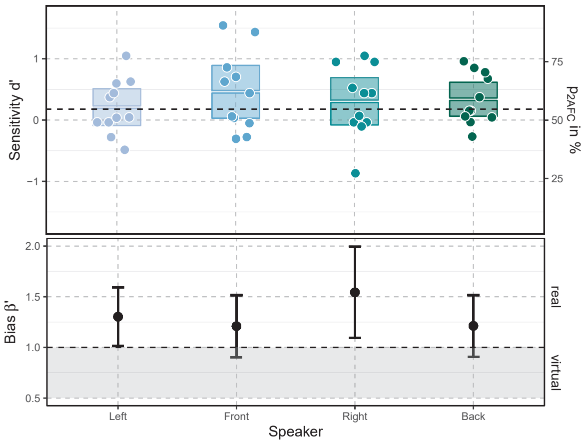

Figure 2. Top: For each speaker separately, the individual sensitivties d' indicated as dots, the mean across subjects d'_{avg} indicated as the white line, and the 95% confidence intervals indicated as the top and bottom of the box. Bottom: mean bias β'_{avg} indicated as a black dot and the 95% confidence intervals for each speaker separately. The dashed line shows the critical d_{min} = 0.177, which corresponds to an P_{AFC} = 0.55%.

due to multiple reflections between the membrane and the head. This could be unnatural for the participants even in the real condition and possibly affect the plausibility results. On the other hand, since the employed HRTFs were measured with a dummy head wearing the AKG K1000, the acoustic effect of the membrane should influence both loudspeaker an virtual reproduction.

As mentioned in Sec.3.4, the choice of an appropriate d_{min} is a point of discussion in the literature. The most stringent criteria of plausibility would be d_{min} = 0. However, this may be hard to achieve. We have used the value d_{min} = 0.177 based on the study by Lindau et al, which however, is a completely arbitrary choice. It is questionable whether it makes any sense in the future to test against a hard plausibility limit or simply to publish the estimated d' and compare them with other studies.

Our results are in line with the results presented by Oberem et al. [18]. The authors reported that with acoustically open and individually equalized headphones binaural reproduction was plausible. Our results showed that even with non-individual HRTFs and non-individual headphone equalization plausible binaural reproduction can be achieved. However, in contrast to Oberem et al. we updated the virtual reproduction not only according to the listener head orientation but also according to translational head movements.

In the present study, we used only the generic HRTF set of one type of dummy head, the Neumann KU100. Therefore, it is questionable how the results can be generalized to different HRTF sets and across subjects. In [26] the influence of HRTF individualization was investigated. It was shown that the influence of best- and worst-matched HRTFs on task performance in VR was rather small. Therefore, it is questionable whether the HRTF set used affects plausibility. Future work is recommended to evaluate the effect of the HRTF set. Furthermore, the effect of listener experience in general is an important factor for the plausibility paradigm. In the present study, we had a limited number of participants to evaluate this. This remains a topic for future research.

6. CONCLUSION

In the present work, we presented the results of a plausibility experiment with a yes/no paradigm that showed that participants could not reliably distinguish between loudspeaker and non-individual binaural headphone reproduction. The experiment was conducted in an anechoic chamber and the virtual sources were synthesized with Neumann KU100 HRTFs, which shows that non-individual HRTFs can produce plausible virtualization of sound sources. The employed Max/MSP rendering pipeline with 6 DoF is also capable of rendering reverberant acoustic spaces. In a follow up series of plausibility experiments, we plan to investigate the influence of listener self translation and the influence of different room acoustics on perceived plausibility.

7. ACKNOWLEDGMENTS

We thank Reality Labs Research at Meta for their financial support of the project.

8. REFERENCES

[1] J. Blauert, *Spatial Hearing*. Camebridge: Hirzel Verlag Stuttgart, 1996.

[2] F. Brinkmann, A. Lindau, M. Vrhovnik, and S. Weinzierl, "Assessing the Authenticity of Individual Dynamic Binaural Synthesis," *EAA Joint Symposium on Auralization and Ambisonics*, vol. 71, no. April, pp. 3–5, 2014.

[3] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[4] A. Lindau and S. Weinzierl, "Assessing the plausibility of virtual acoustic environments," *Acta Acustica united with Acustica*, vol. 98, no. 5, pp. 804–810, 2012.

[5] M. Hofer, T. Hartmann, A. Eden, R. Ratan, and L. Hahn, "The Role of Plausibility in the Experience of Spatial Presence in Virtual Environments," *Frontiers in Virtual Reality*, vol. 1, no. April, pp. 1–9, 2020.

[6] M. Vorländer, *Auralization*. Aachen, Germany: Springer-Verlag Berlin Heidelberg, 1 ed., 2008.

[7] J. Ahrens and C. Andersson, "Perceptual evaluation of headphone auralization of rooms captured with spherical microphone arrays with respect to spaciousness and timbre," *The Journal of the Acoustical Society of America*, vol. 145, no. April, pp. 2783–2794, 2019.

[8] R. Duraiswami, D. N. Zotkin, Z. Li, E. Grassi, N. A. Gumerov, and L. S. Davis, "High Order Spatial Audio Capture and its Binaural Head-Tracked Playback over Headphones with HRTF Cues," in *Proceedings of the 119th AES convention*, (New York, USA), pp. 6540 (pp. 1574–1579), Audio Engineering Society, 2005.

[9] V. Pulkki, "Spatial sound reproduction with directional audio coding," *Journal of the Audio Engineering Society*, vol. 55, no. 6, pp. 503–516, 2007.

[10] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," *Journal of the Audio Engineering Society*, vol. 61, no. 1/2, pp. 17–28, 2013.

[11] A. Neidhardt and A. M. Zerlik, "The Availability of a Hidden Real Reference Affects the Plausibility of Position-Dynamic Auditory AR," *frontiers in Virtual Reality*, vol. 2, no. September, pp. 1–17, 2021.

[12] F. Brinkmann, L. Aspöck, D. Ackermann, S. Lepa, M. Vorländer, and S. Weinzierl, "A round robin on room acoustical simulation and auralization," *The Journal of the Acoustical Society of America*, vol. 145, no. 4, pp. 2746–2760, 2019.

[13] D. Ackermann, F. Fiedler, F. Brinkmann, M. Schneider, and S. Weinzierl, "On the Acoustic Qualities of Dynamic Pseudobinaural Recordings," *Journal of the Audio Engineering Society*, vol. 68, no. 6, pp. 418–427, 2020.

[14] T. Mckenzie, N. Meyer-Kahlen, S. J. Schlecht, and T. Lokki, "Transfer-Plausibility of Binaural Rendering with Different Real-World," in *Proceedings of the 48th DAGA*, (Stuttgard), pp. 1–4, 2022.

[15] S. V. Amengual Garí, J. M. Arend, P. Calamia, and P. W. Robinson, "Optimizations of the Spatial Decomposition Method for Binaural Reproduction (in press)," *Journal of the Audio Engineering Society*, no. December, 2020.

[16] S. Werner, F. Klein, T. Mayenfels, and K. Brandenburg, "A summary on acoustic room divergence and its effect on externalization of auditory events," in *Proceedings of the 8th International Conference on Quality of Multimedia Experience*, no. March 2019, (Lisbon, Portugal), 2016.

[17] J. M. Arend, M. Ramírez, H. R. Liesefeld, and C. Pörschmann, "Do near-field cues enhance the plausibility of non-individual binaural rendering in a dy-

namic multimodal virtual acoustic scene?," *Acta Acustica*, vol. 5, no. 3, 2021.

[18] J. Oberem, B. Masiero, and J. Fels, "Experiments on authenticity and plausibility of binaural reproduction via headphones employing different recording methods," *Applied Acoustics*, vol. 114, pp. 71–78, 2016.

[19] C. Pörschmann, J. M. Arend, and R. Gillioz, "How wearing headgear affects measured head-related transfer functions," in *Proceedings of the EAA Spatial Audio Signal Processing Symposium*, (Paris, France), pp. 49–54, 2019.

[20] B. Bernschütz, "A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100," in *Proceedings of the 39th DAGA*, (Meran), pp. 592—-595, 2013.

[21] J. M. Arend, A. Neidhardt, and C. Pörschmann, "Measurement and Perceptual Evaluation of a Spherical Near-Field HRTF Set," *Tonmeistertagung - VDT International Convention*, no. November, pp. 52–55, 2017.

[22] H. Stanislaw and N. Todorov, "Calculation of signal detection theory measures," *Behavior Research Methods, Instruments, and Computers*, vol. 31, no. 1, pp. 137–149, 1999.

[23] R. D. Morey and J. N. Rouder, "Bayes Factor Approaches for Testing Interval Null Hypotheses," *Psychological Methods*, vol. 16, no. 4, pp. 406–419, 2011.

[24] J. Oberem, J.-G. Richter, D. Setzer, J. Seibold, I. Koch, and J. Fels, "Experiments on localization accuracy with non-individual and individual HRTFs comparing static and dynamic reproduction methods," *bioRxiv*, pp. 702–705, 2018.

[25] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993.

[26] D. Poirier-Quinot and B. F. G. Katz, "Assessing the Impact of Head-Related Transfer Function Individualization on Task Performance: Case of a Virtual Reality Shooter Game," *Journal of the Audio Engineering Society*, vol. 68, no. 4, pp. 248–260, 2020.