PLAUSIBILITY OF SOUND-SOURCE DIRECTIVITY RESYNTHESIS IN A CONTEMPORARY EDISON TEST

Otavio Colella Gomes[1], Nils Meyer-Kahlen[2, 3], Winfried Lachenmayr[4], Tapio Lokki[2]

[1] *Erich Thienhaus Institute, Detmold University of Music, Detmold, Germany*
[2] *Acoustics Lab, Dpt. of Information and Communications Engineering, Aalto University, Aalto, Finland*
[3] *Dyson School of Design Engineering, Imperial College London, UK*
[4] *Werner Genest und Partner Ingenieurgesellschaft mbH, Unterschleißheim b. München, Deutschland,*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Preliminary document

Abstract

At the beginning of the 20th century, the Edison company presented so-called "tone tests" to promote their new phonograph. To demonstrate its quality, listeners were presented with live performances and reproductions of different sound sources. Here, we perform a contemporary Edison test that assesses the influence of different methods for resynthesizing sound-source directivity on the plausibility of musical reproduction in a chamber music hall. Performances of violin, oboe and trumpet and reproductions thereof were played behind an acoustically transparent curtain. The participants then had to decide if it was a live performance or a loudspeaker reproduction. The reproduction methods included a studio loudspeaker and a twelve-channel dodecahedral speaker that was operated either with a one-channel recording fed to all drivers or with a twelve-channel recording from a dodecahedral array surrounding a performing musician. Additionally, beamforming was applied in one condition to enhance low-frequency control. In about 48% of all trials, participants accepted the reproduction as a live performance. There were modest differences between the reproduction methods, indicating that reproductions using the twelve-channel array, where every driver is fed its own signal, might be more plausible than using omnidirectional reproduction or a studio speaker.

Keywords: *plausibility, directivity, sound source*

1. Introduction

The goal of reproducing a sound source such that it is believed to be real is inherent in the history of audio technology since the very beginning. Between 1915 and 1925, thousands of so-called "tone tests" were conducted to promote the new diamond disk phonograph by the Edison company [1]. In these tests, listeners compared sound reproduced using the phonograph to live performances. Newspaper reports from the time are overwhelmingly positive, stating that it was sometimes impossible to tell the difference [1]. Unfortunately, no response data of specific tests is available today. After all, the purpose of these tests was advertisement, not scientific research.

Interestingly, while such tests did not become a standard methodology in the coming decades of audio reproduction development, similar methods are now gaining relevance in the field of audio for augmented reality (AR), with experiments referred to as plausibility tests [2], [3], [4], [5]. A theoretical background of plausibility was developed by Kuhn-Rahloff, who defines it as the agreement of a stimulus with a listener's internal reference [6]. Lindau and Weinzierl proposed a specific test design for assessing plausibility [2]. Therein, either a loudspeaker reproduction or a headphone rendering is presented in each trial, and participants must decide whether the presentation was loudspeaker ("real") or headphone ("virtual") playback.

Here, we employ a similar experimental design in which listeners must decide whether a presented sound originates from a live musician on stage or from a loudspeaker reproduction. This makes the subject and setting similar to the tone test demonstrations conducted a century ago, but implemented under a rigorous scientific protocol.

The aim of our experiment is not only to test the plausibility of loudspeaker reproduction as such, but to evaluate the plausibility of sound reproduction using different directivity resynthesis techniques employing a modified compact dodecahedral loudspeaker array. In Edison's time, reproducing sound with the full

Corresponding author: *otavio.colella@gmail.com.*
Copyright: *©2026 Colella Gomes et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

audio bandwidth, sufficient SNR, and dynamic range were challenging problems. Today, generating clean and dynamical playback with the correct frequency response in front of a loudspeaker is possible. An aspect that is typically still not considered is the reproduction of the correct frequency response in all directions, i.e., the correct directivity. In addition to the variation of the coloration of the sound when listening to reproduction from different directions, directivity influences the way sound is coupled into the room, potentially causing differences in the perception of room acoustic attributes [7], [8].

In concert hall research, in-situ loudspeaker reproductions of anechoic music have nevertheless been widely used, with early concert hall evaluations employing two loudspeakers to represent a full orchestra [9]. Later, a whole loudspeaker orchestra, consisting of many studio loudspeakers, has been used for perceptual evaluation of concert halls [10]. Each one of these speakers does not offer an accurate resynthesis of the directivity of each instrument. However, in [11], the power response of loudspeakers were compared to those of orchestral instruments and loudspeaker arrangements were selected accordingly. For strings, for example, two loudspeakers were used.

More systematic directivity resynthesis has been developed by [12], [13], for example. In these approaches, spherical loudspeaker arrays with individually controllable drivers are employed. For resynthesis, a surrounding recording of a sound source is transformed into the spherical harmonics (SH) domain. Upon playback, the SH signals are decoded by a multiple-input multiple-output (MIMO) filterbank that outputs the loudspeaker driving signals. Instead of using a running multichannel recording, methods can also employ a single-input multiple-output (SIMO) model of directivity [12], in which a directivity is applied to a single-channel recording before decoding. Such an approach was used in [14], for example.

Here, we present a contemporary Edison test that follows a plausibility paradigm and tests the influence of reproduction directivity. Specifically, we test whether a simple resynthesis approach, employing twelve drivers in the loudspeaker array and twelve microphones to record the performance, can yield a more plausible reproduction than that from a studio loudspeaker or an omnidirectional source. We also include a condition in which the low-frequency directivity of the drivers was enhanced using SH-domain beamforming.

2. Methods

2.1. Source Resynthesis

For the test, performances of three instruments with different directivity patterns were used: a violin, a trumpet, and an oboe. The violin is the most omnidirectional instrument in this test. The radiation projects mostly to the front and upper hemisphere with no pronounced additional lobes. At low frequencies (below 1 kHz) the upper and lower hemispheres are more balanced, but it still maintains an emphasis to the front [15]. The trumpet, on the other hand, is the most directional instrument included here. At very low frequencies, it presents an omnidirectional pattern, but as frequency increases, a strong main frontal lobe forms quickly [16]. The oboe presents a main lobe directed forward and slightly downward. This is emphasized at frequencies above 1 kHz; however, at 2 kHz a secondary lobe pointing upwards is also present [15], [17].

The performances of the three instruments were recorded with a surrounding dodecahedral microphone array in the semi-anechoic chamber of Müller-BBM, similar to [18]. To reduce the effect of the reflecting floor, thick absorption panels were placed on the floor below the musician and microphones. The excerpts used in this experiment were:

Violin: 13 s of the Allemande – Partita no. 2 in D minor BWV 1004 by J. S. Bach. Fundamental frequencies ranging from 294 Hz to 932 Hz.

Oboe: 17 s of the first movement *Allegro* from the wind octet serenade KV388 by W. A. Mozart. Fundamental frequencies ranging from 587 Hz to 932 Hz

Trumpet: 14 s of the first movement *Allegro con spirito* of the Trumpet Concerto in E by J. N. Hummel. Fundamental frequencies ranging from 311 Hz to 698 Hz.

During the experiment, different reproduction methods were used to map the signals to the dodecahedral loudspeaker, a modified version of the m|dod 250C from Müller-BBM with a radius of 12.5 cm. In this version, all drivers could be accessed individually. To drive them, the Innosonix MA32-LP multichannel amplifier was used. The first method is direct mapping (DM), in which twelve microphone signals are assigned to each of the twelve loudspeaker drivers. This method is expected to approximate the directional energy distribution at high frequencies, where the loudspeaker drivers are directional by nature, but cause uncontrolled patterns at low frequencies, where the drivers are nearly omnidirectional. As an improvement, the beamforming-enhanced direct mapping (BEDM) approach applies spherical beamforming to increase the directivity of each driver below 1.2 kHz. For the uniform, dodecahedral layout used here for capture and reproduction, the control is equivalent to the mentioned SH directivity control methods developed by Zotter [12]. At high frequencies, it is equivalent to the DM approach.

It is noted that by employing twelve recording positions and twelve drivers, none of the methods are expected to authentically reproduce any target instrument pattern, due to their much higher spatial complexity. However, the difference between a high- and a low-resolution directivity pattern is not always perceptible [19] even in direct comparison, and even more so when the goal is plausibility.

Preliminary document

These methods employing separate signals for all twelve channels of the dodecahedron were compared to two reference cases. These methods only used the signal recorded in front of the musician. For the first reference case, this signal was mapped to all drivers, creating a nearly omnidirectional source. The second reference was a Geithain MO-1 studio loudspeaker.

Furthermore, two different equalization approaches were tested for each dodecahedron reproduction method. In the "diffuse EQ" condition, equalization was based on the mean magnitude spectrum of the surrounding measurements of the loudspeaker array, with the methods reproducing an omnidirectional pattern. Consequently, the diffuse EQ created is independent of the instrument. For this, the directivity of all drivers was measured in the anechoic chamber of the University of Applied Sciences Ostwestfalen-Lippe. High resolution 3D directivity measurements of each driver were performed using the Four Audio ELF system, with the microphone positioned 1.5 m from the center of the driver and a resolution of 5°.

To create the diffuse EQ, the driver responses were added. Then, the magnitude spectra of all measurements surrounding the loudspeaker were averaged, the response was inverted, and a minimum-phase equalizer was designed. Then, this filter was applied to all twelve input channels. The directivity of the studio speaker deviates strongly from omnidirectional. Using diffuse field equalization would lead to an excess of high frequencies. Therefore, an on-axis measurement was used for equalization of the studio speaker.

The "manual EQ" condition was an individual equalizer for each instrument adjusted manually by ear in direct comparison to a real source. This condition was included because coloration was expected to be an important factor contributing to plausibility, and the described instrument-independent diffuse EQ was not expected to fully remove it. The manual equalization was performed by two of the authors in a lecture room, different from the room in which the plausibility experiment took place. Also, here, one global filter was found and applied to all driver signals.

2.2. Listening Experiment

The experiment was conducted in the chamber music hall "Kupferhaus" in Planegg, Germany. The room has a volume of 3000 m^3 and 460 seats that were covered with cloth to simulate a fully occupied audience, bringing the reverberation time to 1.26 s (EDT) and 1.19 s (T_{20}), respectively, at 1 kHz. The sources were positioned at the center of the stage, 2.7 m from its edge. Participants were seated behind an acoustically transparent curtain, at two different positions, one (R1) directly in front of the sources at a distance of 7 m and the second (R2) in the same row of seats but 3.41 m to the side (see Fig. 1). The first half of the test was conducted at one position, the second half at the other. This allowed for two participants to take part at the same time. In total, fifteen listeners (mean age 47.3 years) with musical or acoustical background

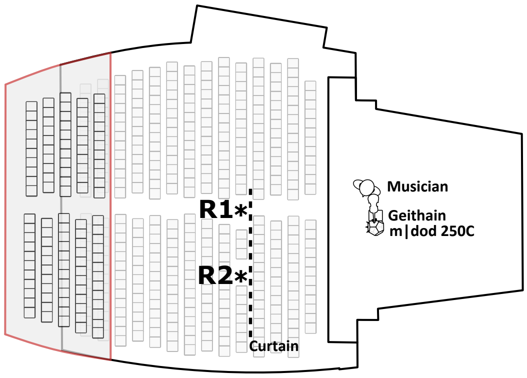


Figure 1: Experiment floorplan at "Kupferhaus".

took part in the test. Most of them were experienced acousticians. Nevertheless, only six of the participants indicated having taken part in a formal listening test within a year before the experiment.

In each trial, either a live performance or one of the reproduction methods was presented. Participants were asked to indicate whether they believed the sound came from a live performance or from a reproduction. The responses were collected using a graphical user interface designed in MATLAB.

The musicians were instructed to direct their instruments in a similar orientation as in the recordings. Three chairs were positioned on the rear of the stage for the musicians to rest during loudspeaker playbacks. A varying number of excerpts before their turn, a musician would quietly walk to the seat closest to the loudspeaker array, get prepared, and wait for the cue from the operator to play their part. This cue was silent and hidden from the test subjects. To prevent that the sound of clicking the mouse would tell which samples were loudspeaker reproductions, a "fake mouse click" was performed before each performance of a musician.

The order of presentation of source reproductions and live performances of all three instruments was randomized for each pair of participants. This was done under the constraint that there could never be the same instrument in two consecutive trials. The purpose of this constraint was to reduce direct comparability between trials. In an ideal plausibility test, the goal is to probe the listeners' internal reference, and the subsequent evaluation relies on the assumption that all trials are mutually independent. For the same reason, no combination of instrument, reproduction method, and seating position occurred more than once. Each participant completed six trials with the live performance and six trials for each of the seven reproduction methods. The design was therefore balanced across the individual source conditions but unbalanced when the trials were grouped into categories of live and reproduced sources, resulting in 6 live and 42 reproduced trials per participant. This structure was chosen to obtain the same number of observations for each reproduction method. Balancing the two categories would have required substantially more live performance trials leading to an increased

Preliminary document

		Given Answer	
		"Musician"	"Reproduction"
True	Musician $N_{\text{Musician}} = 90$	N_{Hits} $= 69$	N_{Misses} $= 21$
	Reproduction $N_{\text{Repr.}} = 630$	$N_{\text{FalseAlarms}}$ $= 301$	$N_{\text{CorrectRejects}}$ $= 329$

Table 1: Pooled response counts across participants and reproduction methods.

risk of inconsistencies between repeated live presentations. Participants were not informed about the relative occurrence of live and reproduced trials. Nevertheless, the imbalance could have influenced their response criterion and is considered when interpreting the acceptance rates, which is discussed below.

3. Results

Table 1 summarizes the responses using signal detection theory (SDT) terminology. Live performances classified as "Musician" were counted as Hits, while those classified as "Reproduction" were Misses. Reproduced trials classified as "Musician" were False Alarms, and correctly identified reproductions were Correct Rejects. Overall, 301 of 630 reproduced trials (47.8%) were accepted as live performances.

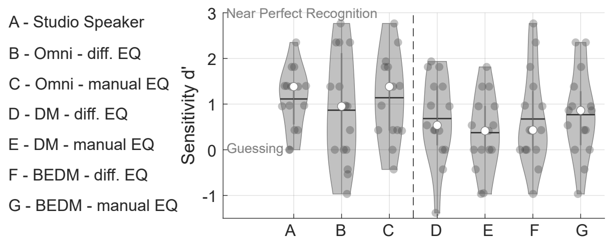
3.1. Sensitivity

The main challenge in analyzing plausibility data is response bias, which refers to the overall tendency to believe that any source is live or not. This tendency is assumed to be different for each participant. For example, it is easy to imagine that naive listeners might be less critical in general than some of the experienced acousticians who participated in this test. Due to such personal bias, looking only at "acceptance as live musician" (i.e., the False Alarm rate), comparison to results from other tests with other participants are not meaningful. To obtain a bias-free measure of plausibility, the sensitivity is computed as

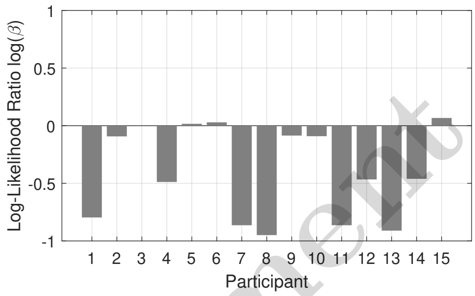
$$d' = \Phi^{-1}(p_{\text{H}}) - \Phi^{-1}(p_{\text{F}}), \qquad (1)$$

where $p_{\text{H}} = N_{\text{Hits}}/N_{\text{Musician}}$ and $p_{\text{F}} = N_{\text{FalseAlarms}}/N_{\text{Repr.}}$, and $\Phi^{-1}(x)$ is the inverse Gaussian cumulative density function.

For the sensitivity analysis, sub-tables similar to Tab. 1 were created for each participant and each reproduction method, in which the first line is identical for all methods, reflecting the responses of each participant to trials with live performances. The second line of each table depends on the method, taking the instruments and positions together. Then, the sensitivity is computed for each participant and method. For SDT analysis, according to Eq. (1), one problem arose. Six out of fifteen participants always recognized the live performances when it was playing (while still accepting reproductions as live on many occasions). This results in an infinite d'. Also, some listeners did not believe some reproduction methods to be live in any trial, with the same effect on d'. To counteract this problem, it is common to introduce a correction



(a) Violinplot of the sensitivity d' for each participant. Black lines show the mean d', the white dot shows the median.



(b) Log-Likelihood ratio. A value < 0 indicates a tendency to respond with "Musician".

Figure 2: Results of SDT analysis.

Table 2: Pairwise comparisons of sensitivity (d'). Reported p-values are uncorrected; an asterisk indicates that the comparison remained significant after Holm correction for multiple comparisons.

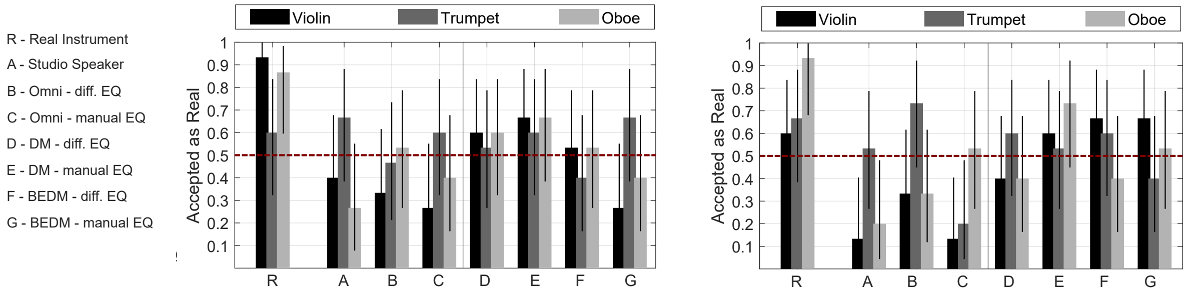
Comparison	$\Delta d'$	$t(14)$	p
Studio vs. DM, diff. EQ	0.43	2.31	.037
Studio vs. BEDM, diff. EQ	0.44	2.39	.032
Studio vs. DM, man. EQ	0.74	3.89	.002*
Studio vs. BEDM, man. EQ	0.34	2.46	.028
Omni vs. DM, diff. EQ	–	–	$> .24$
Omni vs. BEDM, diff. EQ	–	–	$> .24$
Omni vs. DM, man. EQ	0.76	3.55	.003*
Omni vs. BEDM, man. EQ	0.37	1.93	.074

term. Whenever p_{H} is 1, or p_{F} is 0, it is replaced by $\tilde{p}_{\text{H}} = 1/(2N_{\text{Musician}})$ and $\tilde{p}_{\text{F}} = 1 - 1/(2N_{\text{Repr.}})$ respectively [20], where the numbers refer to the total number of trials for live performances and the total occurrences of the reproduction method in the sub-tables $N_{\text{Musician}} = N_{\text{Repr.}} = 6$. The result of the sensitivity analysis is shown in Figure 2a, where $d' = 0$ indicates that participants could not distinguish between musicians and reproductions at all; a value of about $d' = 3$ implies near-perfect recognition. A negative d' implies audible differences that are falsely associated with live performance or reproductions.

Median sensitivities lie between $d' = 0.42$ for direct mapping with manual EQ and $d' = 1.38$ for the studio speaker and omnidirectional speaker with manual EQ. A repeated-measures ANOVA showed a significant effect of reproduction method on d', $F(4.32, 60.49) = 4.00$, $p = .005$, $\eta^2_{\text{G}} = .069$, with Greenhouse-Geisser correction. Pairwise comparisons are summarized in Table 2.

Interestingly, the median sensitivity for all four reproduction methods that take the directivity into account (DM, BEDM, each with diffuse and manual EQ) is

Preliminary document



(a) Central Position (R1) **(b)** Side Position (R2)

Figure 3: Percentage of reproduced trials accepted as live performances. Red dashed line indicated acceptance of 0.5. Black lines indicate 95% confidence intervals of the binomial distribution.

lower than that of the studio speaker and the omnidirectional reproductions. However, after Bonferroni-Holm correction for multiple pairwise comparisons, only two differences remained significant. The DM with manual EQ presented significantly lower sensitivity than both the studio loudspeaker and the omnidirectional loudspeaker with manual EQ. No other pairwise comparison was significant after correction, including the BEDM approach, which is expected to better resynthesize the directivity at low frequencies.

To compute the log-likelihood ratio as an indicator of response bias

$$\log \beta = \frac{1}{2}\left[\left[\Phi^{-1}(p_F)\right]^2 - \left[\Phi^{-1}(p_H)\right]^2\right], \qquad (2)$$

all methods were taken together, so that one subtable was obtained for each participant. The results are seen in Figure 2b, where it becomes clear that the present listeners were generally liberal in their response behavior. Only three log-likelihood ratios are positive; the others indicate a tendency to respond with “Musician”. This bias might not only depend on the credulity of the participants, but also on the fact that since the number of trials was the same for the live performances and each one of the seven reproductions, the overall number of reproduction trials per participant was much larger than the number of live performances (42 vs. 6). Although the participants were not given any information with regards to the distribution, some might have assumed that they should be roughly balanced. A strong bias paired with the low number of musician trials has led to some participants always classifying the musicians as real, making the described correction necessary. This can distort d' estimates. Also, the variance of d' increases when high response bias drives hit or false-alarm rates toward extreme values, where the Φ^{-1} is highly sensitive to small changes in observed proportions.

3.2. Acceptance as live musician

For SDT analysis, instruments and positions had to be taken together. In contrast, Fig. 3 shows the rate of acceptance as a live musician across participants in more detail for different methods, instruments, and positions in the hall, with 95 % confidence intervals of the binomial distribution. As discussed, these acceptance rates cannot be assumed to be unbiased, so their absolute values should be interpreted with care. Furthermore, the limited number of data points for every combination leads to relatively large confidence intervals. Nevertheless, some differences may still be discussed. For example, the studio loudspeaker was more commonly accepted as a live instrument when the trumpet signal was played over the other signals. Especially at the side position, the studio loudspeaker was very rarely accepted as live for the violin and the oboe, but regularly for the trumpet. This result would be easily explained by an influence of directivity on the acceptance rates, as the directional properties of the studio speaker and the trumpet are the closest to each other. Interestingly, in the side position, also the omnidirectional reproduction resulted in a high rating for the trumpet.

As seen in the SDT analysis, a positive trend is observed for the reproduction methods that try to mimic the directivity. Consistently high ratings for all positions and instruments are, for example, obtained for the direct mapping approach with manual equalization. Replacing the manual tuning with the instrument-independent diffuse EQ reduced ratings mainly at the side position for the oboe and the violin.

4. Conclusion

This contemporary Edison test examined whether directivity-informed reproduction yield more plausible reproduction of musical instruments in a concert hall than using a studio loudspeaker or an omnidirectional source.

First of all, we have seen that it is frequently possible to evoke the illusion that a reproduced sound source is in fact a live performance; it was reached in 48% of all trials. When the data was pooled together, SDT analysis showed lower sensitivity for the directivity informed methods, however, after Bonferroni-Holm correction only DM with manual EQ differed significantly from the two reference reproductions and no significant advantage was found for BEDM.

The experiment design was limited by the imbalance between reproduced and live trials (42 vs. 6) and by the small number of observations per condition, which produced strong response biases towards acceptance as real source, and uncertain sensitivity estimates. Future experiments should test fewer reproduction conditions at once, while increasing the number of live

Preliminary document

performances.

5. Acknowledgments

We would like to thank Jithin Thilakan and HfM Detmold for helping with directivity measurements and recordings. The experimental portion of this work was completed while Otavio Colella Gomes was employed at Müller-BBM BSO GmbH and Winfried Lachenmayr at Aalto University and Müller-BBM BSO GmbH. This research was part of the VRACE project that has received funding from the EU's Horizon 2020 research and innovation program under the Marie Skłodowska Curie actions (grant agreement number 812719.)

References

[1] E. Thompson, "Machines, Music, and the Quest for Fidelity: Marketing the Edison Phonograph in America, 1877–1925," *Musical Quarterly*, vol. 79, no. 1, pp. 131–171, 1995. DOI: 10.1093/mq/79.1.131.

[2] A. Lindau and S. Weinzierl, "Assessing the Plausibility of Virtual Acoustic Environments," *Acta Acust*, vol. 98, no. 5, pp. 804–810, Sep. 2012. DOI: 10.3813/AAA.918562.

[3] A. Neidhardt and A. M. Zerlik, "The Availability of a Hidden Real Reference Affects the Plausibility of Position-Dynamic Auditory AR," *Front. Virtual Real.*, 678875th ser., vol. 2, 2021. DOI: 10.3389/frvir.2021.678875.

[4] N. Meyer-Kahlen, S. V. Amengual Garí, S. J. Schlecht, and T. Lokki, "Testing Auditory Illusions in Augmented Reality: Plausibility, Transfer-Plausibility and Authenticity," *J. Audio Eng. Soc.*, vol. 72, no. 11, pp. 797–812, 2024. DOI: 10.17743/jaes.2022.0178.

[5] F. Stärz, S. Van De Par, S. Roßkopf, L. O. H. Kroczek, A. Mühlberger, and M. Blau, "Comparison of binaural auralisations to a real loudspeaker in an audiovisual virtual classroom scenario: Effect of room acoustic simulation, HRTF dataset, and head-mounted display on room acoustic perception," *Acta Acustica*, vol. 9, p. 31, 2025. DOI: 10.1051/aacus/2025012.

[6] C. Kuhn-Rahloff, "Prozesse der Plausibilitätsbeurteilung am Beispiel ausgewählter elektroakustischer Wiedergabesituationen," (in German), Ph.D. dissertation, Universität Berlin, 2011.

[7] F. Otondo, J. H. Rindel, and K. Lyngby, "The Influence of the Directivity of Musical Instruments in a Room," *Acta Acust*, vol. 90, pp. 1178–1184, 2004.

[8] L. M. Wang and M. C. Vigeant, "Evaluations of output from room acoustic computer modeling and auralization due to different sound source directionalities," *Appl. Acoust.*, vol. 69, no. 12, pp. 1281–1293, Dec. 2008. DOI: 10.1016/j.apacoust.2007.09.004.

[9] M. R. Schroeder, D. Gottlob, and K. F. Siebrasse, "Comparative study of European concert halls: Correlation of subjective preference with geometric and acoustic parameters," *J. Acoust. Soc. Am.*, vol. 56, no. 4, pp. 1195–1201, Oct. 1974. DOI: 10.1121/1.1903408.

[10] T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo, "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles," *J. Acoust. Soc. Am.*, vol. 132, no. 5, pp. 3148–3161, Nov. 2012. DOI: 10.1121/1.4756826.

[11] J. Pätynen, "A virtual symphony orchestra for studies on concert hall acoustics," PhD Thesis, Aalto University, Helsinki, 2012.

[12] F. Zotter, "Analysis and Synthesis of Sound-Radiation with Spherical Arrays," PhD Thesis, Universität für Musik und Darstellende Kunst Graz, 2009.

[13] M. Pollow and G. K. Behler, "Variable Directivity for Platonic Sound Sources Based on Spherical Harmonics Optimization," *Acta Acust*, vol. 95, no. 6, pp. 1082–1092, Nov. 2009. DOI: 10.3813/AAA.918240.

[14] M. T. Neal and M. C. Vigeant, "A Compact Spherical Loudspeaker Array for Efficiently Recreating Instrument Directivities," *J. Audio Eng. Soc.*, vol. 68, no. 11, pp. 796–809, Dec. 2020. DOI: 10.17743/jaes.2020.0054.

[15] J. Pätynen and T. Lokki, "Directivities of Symphony Orchestra Instruments," *Acta Acust*, vol. 96, no. 1, pp. 138–167, Jan. 2010. DOI: 10.3813/AAA.918265.

[16] A. C. Marruffo, J. Thilakan, A. Hofmann, V. Chatziioannou, and M. Kob, "High-resolution 3D directivity measurements of a trumpet," in *DAGA*, 2022.

[17] J. Meyer, *Acoustics and the Performance of Music*, 5th ed. New York: Springer Science & Business Media, 2009.

[18] O. C. Gomes, W. Lachenmayr, J. Thilakan, and M. Kob, "Anechoic Multi-Channel Recordings of individual String Quartet musicians," in *Immersive and 3D Audio*, Sep. 2021. DOI: 10.1109/I3DA48870.2021.9610912.

[19] A. Corcuera-Marruffo, N. Meyer-Kahlen, and T. Lokki, "Audibility of reduced spatial resolution in musical instrument directivity," *Acta Acustica*, vol. 9, May 2025. DOI: 10.1051/aacus/2025016.

[20] N. A. Macmillan and C. D. Creelman, *Detection theory: a user's guide*, 2nd. Mahwah, N.J: Lawrence Erlbaum Associates, 2005, ISBN: 978-0-8058-4230-2 978-0-8058-4231-9.

Preliminary document.