Audio Engineering Society

Convention Express Paper 350

Presented at the 158th Convention
2025 May 22–24, Warsaw, Poland

This Express Paper was selected on the basis of a submitted synopsis that has been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This Express Paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Perceptual evaluation of professional point and line sources for immersive audio applications

Leon Merkel[1], Franz Zotter[1,2], and Philip Coleman[3]

[1]*University of Music and Performing Arts, Graz, Austria*
[2]*Institute of Electronic Music and Acoustics, Graz, Austria*
[3]*L-Acoustics, 67 Southwood Lane, London N6 5EG*

Correspondence should be addressed to Leon Merkel (leon.merkel@gmail.com)

ABSTRACT

Immersive sound reinforcement aims to create a balanced perception of sounds arriving from different directions, establishing an impression of envelopment over the audience area. Current perceptual research shows that coverage designs featuring nearly constant decay (0 dB per distance doubling) preserve the level balance among audio objects in the mix. In contrast, a -3 dB decay supports a more uniform sensation of envelopment, especially for off-center listening positions. For practical reasons, point-source loudspeakers remain widely used for immersive audio playback in mid-sized venues. However, their -6 dB level decay per distance doubling can conflict with the design goals outlined above. This paper investigates how much perceived difference there is between point-source and line-source setups, using eight surrounding loudspeakers side-by-side covering a $10\,\mathrm{m} \times 7\,\mathrm{m}$ audience area. The perceptual qualities object level balance, spatial definition, and envelopment were compared in a MuSHRA listening experiment, and to explain the results, acoustic measurements were carried out to capture room impulse responses (RIRs) and binaural room impulse responses (BRIRs). To verify that the perceptual results can be reproduced on headphones, BRIRs were used in a follow-up experiment. Both the loudspeaker and headphone-based experiments confirm the expectation from earlier experiments, but now with dedicated, professional state-of-the-art loudspeakers.

1 Introduction

The emergence of immersive mixing technologies for large-scale sound reinforcement allows audiences to experience balanced imaging of direct sound objects on the stage. Further, it increases immersion by placing the audience within the audio scene instead of solely in front of it [1]. Exemplarily, Corteel et al. [2] presented particular loudspeaker deployments and panning algorithms for large-scale deployments.

The system and coverage design for immersive deployments have recently been investigated. Gölles et al. [3] evaluated coverage design goals with small-size DIY

3D-printed line arrays [4] of adjustable level decay per doubling of distance (/dod). They achieved 0 dB/dod by the geometric curving of the arrays, and −3 dB/dod or −6 dB/dod by applying appropriate delays. With these arrays in a room, −6 dB/dod was always least preferred, −3 dB/dod reached the most uniform perceived diffuse envelopment, and a constant level with 0 dB/dod was found to preserve a uniform perceived sound-object level balance. These observations were in line with laboratory experiments in a nearly anechoic room using gains and delays to simulate the loudspeaker throw [5, 6, 7]. For practical deployments, Gölles et al. [8] found that direct-sound decay ranging between 0 dB and −3 dB/dod is often perceptually satisfying as a compromise of both ideal conditions for enveloping sounds and ideal conditions for a mixture of direct-sound objects.

The preceding experiments have covered: (i) anechoic direct-sound directions and level decays under laboratory conditions, and (ii) DIY line arrays in a room, with delays profiles configured for different level decays. This paper addresses both of these aspects, aiming to answer the practical question: *how do professional point-source loudspeaker systems and professional line-source loudspeaker systems in the same room compare with each other, side-by-side?*

In section 2, the experimental setup and design of the listening experiments are illustrated, and their results are presented in section 3. Correlations with acoustic metrics are explored in section 4, and section 5 concludes the findings.

2 Experimental Setup and Design

In this section we introduce the physical loudspeaker deployment, the test protocol and a headphone-based experiment mirroring the on-site listening tests.

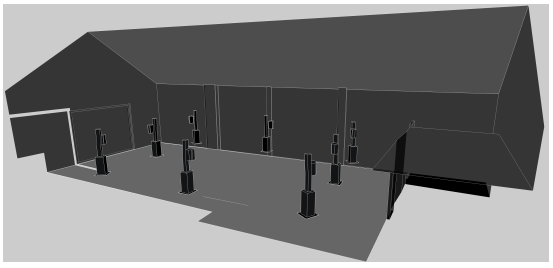


Fig. 1: Model of L-Acoustics Immersive Lab.

2.1 System configurations

The first listening experiment was conducted at L-Acoustics Immersive Lab, cf. Figure 1, with a mean reverberation time of 0.8 s (100 Hz –10 kHz). An eight-channel surround loudspeaker deployment with 45 ° angular resolution covers an audience area of 10 m × 7 m, as shown in Figure 2. The line source used was the compact, fixed curvature Syva[1] loudspeaker, which reached an average free-field A-weighted direct-sound decay of −2 dB/dod in the distance range of 1–8 m[2] with a placement that vertically centered Syva's waveguide section at a height of 2.1 m. For comparison, the coaxial point source X8[3] vertically centered at the height of 2.1 m achieved a A-weighted direct-sound decay of −5 dB/dod.

To minimize the chance of identifying the loudspeaker type during the experiment, both units were placed as close to each other as possible. Level inequalities due to the distances of the sources to the origin were compensated and the frequency responses of all loudspeaker around the origin are equalized to ensure a flat magnitude from 125 Hz to 16 kHz. All loudspeakers were measured at nine locations around the origin and the location average of the 1/3-octave band smoothed magnitudes were equalized with bespoke FIR-filters (HF: zero-phase, LF: minimum-phase).

The tested configurations are summarized in Table 1. Three idealized references I1, I2, I3 are position-specifically equalized (*posEQ* in Table 1), ensuring

[1] https://www.l-acoustics.com/products/syva/

[2] Computed from Soundvision, https://www.l-acoustics.com/products/soundvision/

[3] https://www.l-acoustics.com/products/X8/

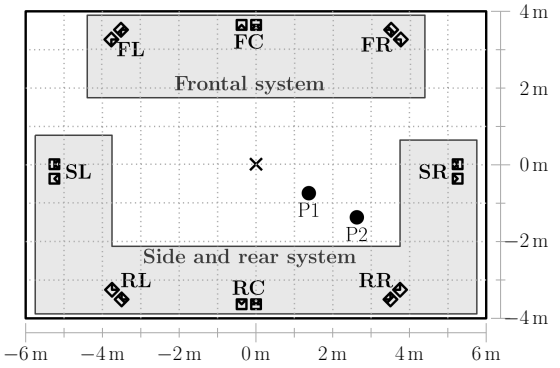


Fig. 2: Loudspeaker deployment for the experiment.

Configurations		Frontal	Side and rear
Idealized	I1	Syva (*posEQ*)	Syva (*posEQ*)
	I2	Syva (*posEQ*)	X8 (*posEQ*)
	I3	X8 (*posEQ*)	X8 (*posEQ*)
Realistic	R1	Syva	Syva
	R2	Syva	X8
	R3	X8	X8
Anchor	A3	X8 (C-9 dB)	X8

Table 1: Tested loudspeaker configurations.

uniform level and magnitude for all loudspeakers at the respective listening position. I1 consists solely of eight Syva line sources, I2 features three Syvas in the frontal positions (FL, FC, FR) and X8 point sources in the side and rear positions (SR, RR, RC, RL, SL) imitating typical immersive setups where frontal line arrays are deployed above the stage, while I3 employs X8s for all directions. Realistic conditions R1, R2, R3, won't achieve simultaneous position-specific corrections, in general, and were equalized for a central listening position. While they differ in equalization, the realistic configurations otherwise mirror the idealized ones, with R1 comprising only Syvas, R2 as a mixed setup with Syvas in front and X8s surrounding, and R3 using solely X8s. The anchor configuration A3 is imbalanced on purpose, by attenuating the center by -9 dB, when employing X8s in all directions.

2.2 Listening experiment design

Participants were asked to evaluate the seven loudspeaker configurations for three spatial qualities at two off-center listening positions, P1 and P2, using a Multi-Stimulus test with Hidden Reference and Anchor (MuSHRA). The user interface[4] randomized the order of presented trials, criteria and setups, allowing reliable judgments on a scale of 0-100 for the spatial qualities against its similarity to the given reference.

2.2.1 Rating Criteria

The three rating criteria are based on the spatial audio quality inventory (SAQI) [9]. Some qualities are assumed to correlate and are summarized into adjusted qualities, including:

- the **object level balance** (OLB), which describes loudness relationships between panned direct sound objects/instruments,
- the **spatial definition** (SD), refers to clarity, distance impression, and direct-to-reverberation ratio,
- the **spatial envelopment** (SE), specifies the sensation of being equally surrounded by sound.

[4] https://git.iem.at/rudrich/mushra

2.2.2 Stimuli

Two musical excerpts of different genres were presented to rate the OLB and SD. The first is a live recording of a Jazz/Pop band, giving distinct instrument differentiation, localizability and transient components. The direct sound of the piano is assigned at FL, the bass guitar and vocals are mixed at FC, and the drums are allocated at FR, aiming for a typical stage while avoiding spatial artifacts from panning. Using *Waves IR-1* convolution reverb plugin, the altered RIRs (without direct sound) from the Kirishima Concert Hall from independent locations provide uncorrelated reverb ($T_{60} = 0.9$ s) with a realistic reflection behavior.

The second excerpt features an 18-piece chamber string orchestra, giving rich harmonic content, high dynamic range and less precise localization of instruments. The nine recorded spot-mics are allocated to front loudspeakers to recreate an orchestra seating arrangement. As already described above convolution reverb ($T_{60} = 1.9$ s) from surround loudspeakers enhance the listening experience with an appropriate acoustic space.

Uncorrelated pink noise from all directions facilitates the assessment of the uniformity of SE. Riedel et al. [5] quantified the ideal direct sound decay for envelopment as -3 dB/dod, therefore an additional gain

$$g_{\mathrm{ls,pos}}^{-3\,\mathrm{dB}} = \sqrt{\frac{d_{\mathrm{ls,origin}}}{d_{\mathrm{ls,pos}}}}, \tag{1}$$

is added to idealize the configurations I1, I2, and I3 for this criterion, where $d_{\mathrm{ls,origin}}$ and $d_{\mathrm{ls,pos}}$ are the distances between loudspeakers and the origin or listening positions, respectively.

2.3 Headphone Experiment

To verify whether the key perceptual outcomes from the on-site experiment translate consistently into a headphone-based scenario, the acoustic paths of all loudspeakers to both listening positions P1 and P2 and the origin were captured with BRIRs. Those allow for creating a virtual framework for the experiment with headphones.

The diffuse field equalization compensates for unwanted effects from the ear canal simulator of the KEMAR dummy head. It filters the BRIRs towards a flat reference from an omnidirectional microphone, enabling realistic-sounding binaural playback through headphones and the individual ear canal. The magnitude of the filter is defined with

$$|H_{\mathrm{eq}}[\omega]| = \sqrt{\frac{\sum^{\mathrm{ls}} |H_{\mathrm{ls}}^{\mathrm{omni}}[\omega]|^2}{\sum^{\mathrm{ls}} \sum^{\mathrm{ear}} |H_{\mathrm{ls,ear}}^{\mathrm{head}}[\omega]|^2}}, \qquad (2)$$

with the on-site measured frequency responses from all utilized loudspeakers to the omnidirectional microphone $H_{\mathrm{ls}}^{\mathrm{omni}}[\omega]$ and the frequency response to both ears of the dummy head $H_{\mathrm{ls,ear}}^{\mathrm{head}}[\omega]$. ω denotes the angular frequency.

For binaural playback in the second experiment, the 3D-printed ultralight circumaural open headphones from [10] were chosen to ensure a high degree of externalization for realistic binaural sound field reproduction. Due to the open construction, extra equalization was appplied to flatten the magnitude of the headphones.

3 Experiment results

In the following, we introduce the results for the on-site and headphone experiments. In the plots, the dots indicate the median and the vertical lines represent the 95% confidence intervals. Using the Wilcoxon sign-rank test with Bonferroni-Holm correction indicate significant difference for values of $p < 0.05$.

3.1 Loudspeaker reproduction

In total, 17 listeners with an average age of 34 years and relevant experience in immersive audio playback participated in the on-site experiment, which took an average duration of 40 min. Two participants were excluded, as the hidden reference was rated lower than 90 in >15% of the test cases. The experimental design is validated, as most participants detected the hidden reference and anchor for all criteria.

3.1.1 Object Level Balance

We observe that idealized and realistic configurations with frontal Syva achieve significantly higher OLB ratings than solely point-source designs, cf. Figure 3, with $p < 0.01$ for both positions P1, and P2 (cf. Table 2). When comparing R1 with R2, no significant advantage ($p > 0.8$) emerges for realistic configurations with Syva loudspeakers in side/rear use at P1 and P2.

3.1.2 Spatial Definition

Concerning SD ratings in Figure 4, frontal Syva consistently outperforms frontal X8 significantly for both the idealized and realistic configurations with $p < 0.03$ (cf. Table 3) at both positions P1 and P2. Similar to the OLB, no significant benefit appears for pure Syva configurations (R1) compared to mixed Syva/X8 configurations (R2) with $p > 0.5$ at P1 and P2. The significantly ($p < 0.01$) lower ratings of I3, compared I1, underlines how point-source directivity more strongly excites the reverberant field, reducing spatial definition.

3.1.3 Spatial Envelopment

The SE ratings in Figure 5 show that both, realistic and idealized setups with Syva in rear/side positions significantly enhance the uniformity of envelopment with $p < 0.03$ for P1 and P2 (cf. Table 4). Comparing R2 with R3, significant differences only occur at P1 ($p = 0.02$).

3.2 Headphone reproduction

In total, 18 listeners with an average age of 29 years and experience in audio engineering participated in the headphone experiment, which took an average duration of 42 min. Similar to the on-site experiment, the experimental design was confirmed, as most participants found the hidden reference and anchor in 85% of the cases (except two participants, who were removed).

p	I1 vs I3	R1 vs R2	R1 vs R3	R2 vs R3
P1	$<$ **0.01**	0.97	$<$ **0.01**	$<$ **0.01**
P2	$<$ **0.01**	0.82	**0.04**	0.12

Table 2: Pairwise comparison of on-site OLB ratings.

p	I1 vs I3	R1 vs R2	R1 vs R3	R2 vs R3
P1	$<$ **0.01**	0.86	$<$ **0.01**	**0.03**
P2	$<$ **0.01**	0.55	**0.02**	0.05

Table 3: Pairwise comparison of on-site SD ratings.

p	I1 vs I2	R1 vs R2	R1 vs R3	R2 vs R3
P1	$<$ **0.01**	$<$ **0.01**	$<$ **0.01**	**0.02**
P2	$<$ **0.01**	**0.02**	$<$ **0.01**	0.53

Table 4: Pairwise comparison of on-site SE ratings.

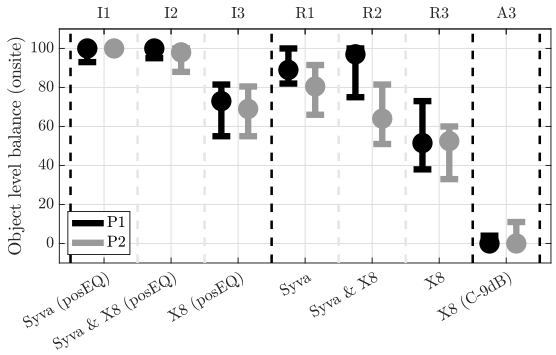


Fig. 3: On-site experiment ratings for the OLB.

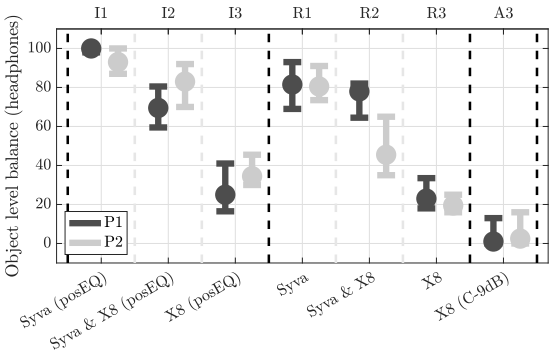


Fig. 6: Headphone experiment ratings for the OLB.

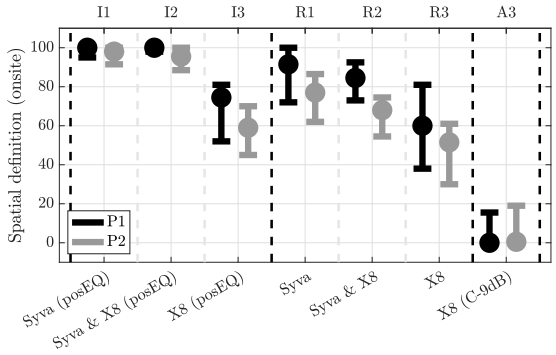


Fig. 4: On-site experiment ratings for the SD.

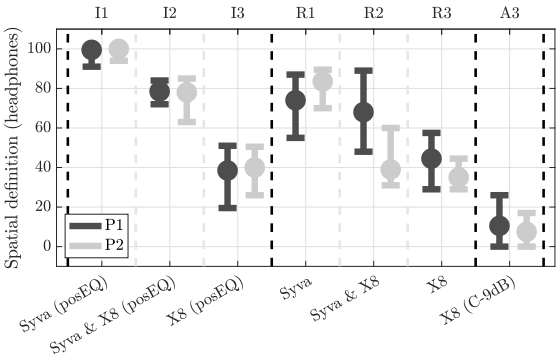


Fig. 7: Headphone experiment ratings for the SD.

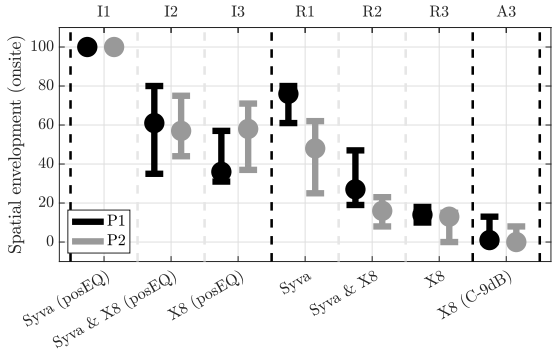


Fig. 5: On-site experiment ratings for the SE.

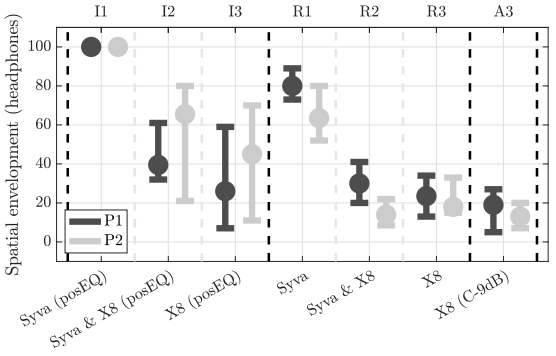


Fig. 8: Headphone experiment ratings for the SE.

3.2.1 Object Level Balance

The binaural reproduction of the experiment follows the on-site experiments trend, cf. Figure 6, for which configurations with Syva loudspeaker in the front again rated significantly better in preserving the object level balance than point-source setups with $p < 0.01$ for both positions P1 and P2 (cf. Table 5). Concerning the OLB, no significant benefit appears for side/rear Syva loudspeaker when comparing R1 with R2 at P1 ($p = 0.58$). At P2 a significant difference between R1 and R2 is observed with $p < 0.01$, while at P1, no significant differences are determined for this configuration pair. All layouts comprising X8 loudspeakers are rated lower than the on-site results. The correlation coefficient of OLB ratings on-site vs headphones is $\mathbf{R}_{\mathrm{OLB}} = \mathbf{0.89}$.

p	I1 vs I3	R1 vs R2	R1 vs R3	R2 vs R3
P1	$< \mathbf{0.01}$	0.58	$< \mathbf{0.01}$	$< \mathbf{0.01}$
P2	$< \mathbf{0.01}$	$< \mathbf{0.01}$	$< \mathbf{0.01}$	$< \mathbf{0.01}$

Table 5: Pairwise comparison of headphone OLB.

p	I1 vs I3	R1 vs R2	R1 vs R3	R2 vs R3
P1	$< \mathbf{0.01}$	0.82	$\mathbf{0.04}$	0.12
P2	$< \mathbf{0.01}$	$< \mathbf{0.01}$	$< \mathbf{0.01}$	1.18

Table 6: Pairwise comparison of headphone SD.

p	I1 vs I2	R1 vs R2	R1 vs R3	R2 vs R3
P1	$< \mathbf{0.01}$	$< \mathbf{0.01}$	$< \mathbf{0.01}$	0.62
P2	$< \mathbf{0.01}$	$< \mathbf{0.01}$	$< \mathbf{0.01}$	0.58

Table 7: Pairwise comparison of headphone SE.

3.2.2 Spatial Definition

The headphone experiment SD ratings, shown in Figure 7, also mirror the on-site experiment observations. Setups with frontally deployed Syva score higher than frontal X8 designs. Exceptionally, the comparison of R2 and R3 show no significant differences with $p > 0.1$ for P1 and P2 (cf. Table 6). In the realistic scenario at P1, no significant benefit emerges from fully Syva-based vs mixed configurations with X8 for side/rear usage with $p = 0.82$. At the same time, P2 shows significant advantages having $p < 0.01$. The lower ratings for X8 configurations indicate that the impact of source directivity on SD as observed in the on-site experiment translates to binaural reproduction. The on-site and headphone ratings for SD correlate with $\mathbf{R}_{\mathrm{SD}} = \mathbf{0.90}$.

3.2.3 Spatial Envelopment

In Figure 8, the SE ratings from the headphone experiment confirm the on-site experiment observation that rear/side Syva significantly improve the uniformity of envelopment for the idealized and realistic setups and both positions P1 and P2 with $p < 0.01$ (cf. Table 7). Compared to the on-site experiment, the SE ratings for the realistic configurations are slightly higher, possibly attributable to the higher amount of perceived reverberation in static binaural playback. The overall rank order of the headphone ratings aligns closely with the on-site experiment exhibiting a correlation of $\mathbf{R}_{\mathrm{SE}} = \mathbf{0.94}$.

3.3 Discussion of experiment results

The on-site ratings highlight that the utilized line-source deployments, although not perfectly matching the ideal 0 dB or -3 dB/dod coverage targets, still preserve the OLB more effectively than point-source configurations. It was shown that frontal line sources primarily influence the OLB due to their decay behaviour, they increase the SD due to their better coverage and higher directivity, hence reduced reverberant field. The sensation of uniform envelopment is enhanced with surrounding line sources. Presumably the colouration of room reflections partly led to poorer ratings of the idealized configuration with point sources when compared to realistic layouts consisting solely of line sources.

It was not studied here but reported that static and non-individual binaural reproduction increases perceived reverberation, which generally reduces OLB and SD ratings. Point sources, already exciting the reverberant field more strongly, yield a pronounced loss in clarity and definition. Additionally, pink noise may shift focus toward tonal coloration rather than the intended spatial criteria. Still, strong correlations ($R > 0.89$) confirm the transferability of rank-order preferences between on-site and headphone experiments.

4 Acoustic Metrics

To model the perceptual ratings, three acoustic metrics were derived from measured RIRs and BRIRs at the listening positions. The front-to-all ratio (FAR), direct-to-diffuse ratio (DDR), and interaural level difference (ILD) are chosen and tailored to reflect the three rating criteria of object level balance, spatial definition, and spatial envelopment, respectively, for pink $1/\sqrt{\omega}$ input signal spectra. To map the logarithmic metrics to the perceptual ratings (0–100), each metric is linearly regressed onto the corresponding experimental ratings. Exemplarily, for FAR the regression to scaled values $\widetilde{\mathrm{FAR}}$ is $\widetilde{\mathrm{FAR}} = \alpha_{\mathrm{FAR}}\,\mathrm{FAR} + \beta_{\mathrm{FAR}}$, where α_{FAR} and β_{FAR} are the optimal regression parameters obtained for both positions and qualities separately by finding the least-squares solution of the metrics with the median ratings of the realistic configurations.

4.1 Front-to-All ratio

The front-to-all ratio (FAR) compares the weakest A-weighted direct sound level from the three frontal loudspeakers to the strongest direct sound level from any

loudspeaker in the system. This reflects the perceptual balance of the mix, as the weakest frontal object must remain clearly audible and not be masked by louder contributions from other directions. The FAR reads as

$$\mathrm{FAR} = 10\log_{10}\left(\frac{\min(\sum^{\omega}\frac{1}{\omega}\,|H_{\mathrm{A}}[\omega]\,H^{\mathrm{dir}}_{\mathrm{front}}[\omega]|^2)}{\max(\sum^{\omega}\frac{1}{\omega}\,|H_{\mathrm{A}}[\omega]\,H^{\mathrm{dir}}_{\mathrm{all}}[\omega]|^2)}\right), \tag{3}$$

where $H_{\mathrm{A}}[\omega]$ is the transfer function of the A-weighting filter, $H^{\mathrm{dir}}_{\mathrm{front}}[\omega]$ transfer functions of the direct sound from the three frontal loudspeakers, and $H^{\mathrm{dir}}_{\mathrm{all}}[\omega]$ the transfer functions of direct sound from all loudspeakers. The FAR metric correlates well with the OLB ratings (Figure 9), showing clear differences between P1 and P2, except for R2, and yielding $\mathbf{R}_{\mathrm{FAR}} = \mathbf{0.97}$ with all values inside the 95% confidence intervals.

4.2 Direct-to-Diffuse ratio

The direct-to-diffuse ratio (DDR) compares the mean A-weighted direct sound energy from the three frontal loudspeakers to the mean diffuse sound energy from the side/rear loudspeakers. By isolating only the early direct sound (first 2.9 ms) and the later diffuse parts (after 2.9 ms) of the RIRs, the metric compares the perceptual balance between frontal directness/clarity and surrounding diffuseness. It reads as:

$$\mathrm{DDR} = 10\log_{10}\left(\frac{\frac{1}{N_{\mathrm{ls,fr}}}\sum^{\mathrm{ls,fr}}\sum^{\omega}\frac{1}{\omega}\,|H_{\mathrm{A}}[\omega]\,H^{\mathrm{dir}}_{\mathrm{fr}}[\omega]|^2}{\frac{1}{N_{\mathrm{ls,su}}}\sum^{\mathrm{ls,su}}\sum^{\omega}\frac{1}{\omega}\,|H_{\mathrm{A}}[\omega]\,H^{\mathrm{dif}}_{\mathrm{su}}[\omega]|^2}\right), \tag{4}$$

with the number of frontal $N_{\mathrm{ls,fr}}$, and side and rear loudspeaker $N_{\mathrm{ls,su}}$. The transfer function $H^{\mathrm{dir}}_{\mathrm{fr}}[\omega]$ is calculated from the RIRs from the frontal sources (direct part of RIR), and $H^{\mathrm{dif}}_{\mathrm{su}}[\omega]$ is the transfer function of the side and rear loudspeaker (diffuse part of RIR). The DDR correlates with SD ratings (Figure 10) at $\mathbf{R}_{\mathrm{DDR}} = \mathbf{0.96}$, reproducing R1 and R2 well, though R3 and A3 at P2 fall just within the interval.

4.3 Interaural Level Difference

The interaural level difference (ILD) gives a good measure of the uniformity of the surrounding sound field, which suits as a comparable metric to the SE. The mean A-weighted level at the left ear is compared to the mean A-weighted level at the right ear, yielding:

$$\mathrm{ILD} = 10\log_{10}\left(\frac{\sum^{\mathrm{ls}}\sum^{\omega}\frac{1}{\omega}\,|H_{\mathrm{A}}[\omega]\,H^{\mathrm{all}}_{\mathrm{left\,ear}}[\omega]|^2}{\sum^{\mathrm{ls}}\sum^{\omega}\frac{1}{\omega}\,|H_{\mathrm{A}}[\omega]\,H^{\mathrm{all}}_{\mathrm{right\,ear}}[\omega]|^2}\right), \tag{5}$$

where $H^{\mathrm{all}}_{\mathrm{left\,ear}}[\omega]$ and $H^{\mathrm{all}}_{\mathrm{right\,ear}}[\omega]$ are the transfer functions of the each ear, respectively. The ILD regression to SE ratings (Figure 11) reaches a correlation of $\mathbf{R}_{\mathrm{ILD}} = \mathbf{0.98}$, matching the ratings across both positions without notable deviation.

4.4 Discussion of acoustic metrics

The metrics introduced in section 4 are linearly regressed to the on-site experiment ratings for each position separately, as no direct comparisons were performed between them. Since the idealized configurations serve no practical purpose, only the realistic setups are included in the regression. Overall, the instrumental metrics FAR, DDR, and ILD show strong

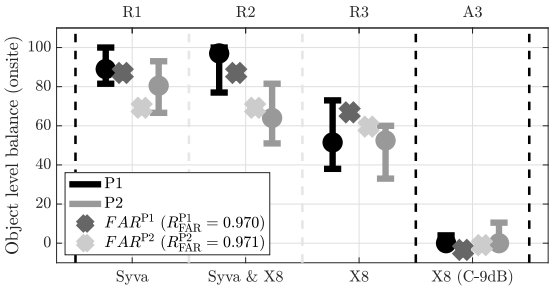


Fig. 9: Regression of FAR to OLB ratings.

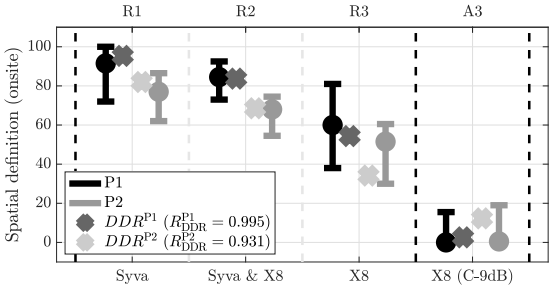


Fig. 10: Regression of DDR to SD ratings.

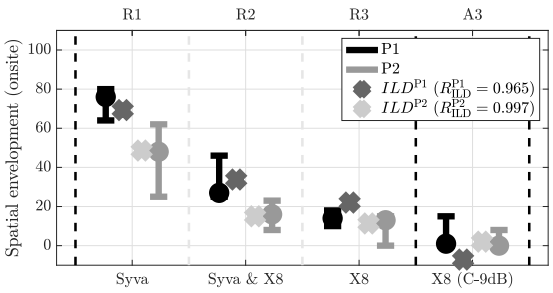


Fig. 11: Regression of ILD to SE ratings.

correlations ($R > 0.93$) with the perceptual data and reliably reflect the behavior of the tested line and point sources. While promising for generalization to other setups, this has not yet been investigated.

5 Conclusion

This work compared professional coaxial point sources and compact fixed-curvature line sources in immersive sound reinforcement, focusing on off-center seats. The line sources were shown to extend the sweet area in which object-level balance is preserved, confirming earlier prototype-based findings [8]. Frontally deployed line arrays were shown to reduce reverberant energy and improve spatial definition, while surrounding line sources were shown to enhance uniform envelopment by approximating the -3 dB/dod design recommendation [3, 5, 6]. By contrast, point-source directivity excites the reverberant field more strongly, lowering all criteria—even in idealized setups. Ratings for headphone-based playback strongly correlated with on-site results ($R > 0.89$), though static and non-individualized binaural reproduction yielded slightly lower object-level balance and definition and higher envelopment ratings. Finally, linear regression of the instrumental metrics (FAR, DDR, and ILD) to experiment ratings achieved correlations $R > 0.93$, underscoring their predictive power for system design and layout decisions.

The preservation of relative mappings in the non-individual, virtual binaural representation is encouraging for evaluating large-scale systems in virtual, headphone-based experiments. This enables perceptual studies of adjustable magnitude and decay behavior, even for specialized deployments. Adapting the presented metrics to generalized immersive layouts with varying source and design types may further improve their predictive accuracy.

Supplementary Data

The measured RIRs, BRIRs at the listening positions and the anonymized experiment ratings from both experiments can be made available upon request.

Acknowledgment

The authors gratefully thank all listeners taking part in the listening experiments. Thanks to L-Acoustics for providing their Immersive Lab for the on-site experiment. The authors are grateful to the Audio Experience Design group from Imperial College for lending their KEMAR dummy head.

References

[1] Paterson, J. and Lee, H., *3D Audio*, Perspective on Music Production, Routledge, 2021, ISBN 9781138590069, doi:10.4324/9780429491214.

[2] L-Acoustics, "Immersive Hyperreal Sound System Design," 2025, accessed: 2025-01-16.

[3] Gölles, L., Frank, M., and Zotter, F., "Improving Surround Sound Reinforcement at Off-Center Listening Positions with Miniature Line Arrays," in *AES Convention: Europe 2024, Madrid*, AES, 2024.

[4] Gölles, L., Zotter, F., and Merkel, L., "Miniature line array for immersive sound reinforcement," in *AES Conference: International Conference on Spatial and Immersive Audio 2023*, AES, 2023.

[5] Riedel, S. and Zotter, F., "Surrounding line sources optimally reproduce diffuse envelopment at off-center listening positions," *JASA Express Letters*, 2(9), 2022, ISSN 2691-1191.

[6] Riedel, S., Frank, M., Zotter, F., and Sazdov, R., "A Study on Loudspeaker SPL Decays for Envelopment and Engulfment across an Extended Audience," *Journal of the AES*, 2024.

[7] Zotter, F., Riedel, S., Gölles, L., and Frank, M., "Acceptable imbalance of sound-object levels for off-center listeners in immersive sound reinforcement," *Fortschritte der Akustik, DAGA, Hamburg*, 2023.

[8] Gölles, L., Frank, M., and Zotter, F., "Evaluating a Dual-Target Line-Array Design for Medium-Scale Surround Sound Reinforcement," in *AES Convention: Europe 2024, Madrid*, AES, 2024.

[9] Lindau, A., Erbes, V., Lepa, S., Maempel, H.-J., Brinkmann, F., and Weinzierl, S., "A Spatial Audio Quality Inventory (SAQI)," 2014, doi:10.14279/DEPOSITONCE-15275.

[10] Mülleder, A. and Zotter, F., "Ultralight Circumaural Open Headphones," in *Proceedings of the 154th AES Convention*, 2023.