Enhancement of the acoustic design workflow through VR technology: the case study of conference room in a historical building

Davide Cetani[a]* | Louena Shtrepi[a] | Angela Guastamacchia[a] | Arianna Astolfi[a]

[a] Dipartimento di Energia,
Politecnico di Torino,
Corso Duca degli Abruzzi, 24, 10129 Torino
* Corresponding author:
davide.cetani@polito.it

Ricevuto: 28/2/2025
Accettato: 17/9/2025

DOI: 10.3280/ria2-2025oa19529
ISSNe: 2385-2615

Speech intelligibility is a key factor in conference room acoustics, affecting the listener's experience and communication effectiveness. Traditional acoustic design relies on objective parameters, often neglecting perceptual validation of both acoustic and aesthetic features. This study explores how virtual reality technologies can support the acoustic design process. Two acoustic treatment solutions for the conference room of the Museo Egizio di Torino were compared: a (1) seamless plaster-based system and (2) parametric felt baffles. Geometrical Acoustic software Ramsete was used to predict their effectiveness, and to generate third-order Ambisonics auralisations. A photogrammetric 3D model was used for visual renderings. Both were used for speech intelligibility tests conducted in the Audio Space Lab of Politecnico di Torino, a small listening environment equipped with a High-Order Ambisonics sound reproduction system and an Oculus Meta Quest 2. Listeners were also asked to express their preference on the visual scenarios obtained with the two treatment solutions. The results showed a preference for the second solution, despite no significant acoustic differences.

Keywords: audiovisual integration, auralisation, immersive experience, speech intelligibility, design tool

Miglioramento del workflow di progettazione acustica attraverso le tecnologie VR: il caso studio di una sala conferenza in un edificio storico

L'intelligibilità del parlato è un aspetto chiave dell'acustica delle sale conferenza, poiché influisce sulla comunicazione. La progettazione acustica normalmente si basa su parametri oggettivi, trascurando la validazione percettiva dell'acustica e dell'estetica. Questo studio indaga l'uso della realtà virtuale a supporto della progettazione acustica. Sono state confrontate due soluzioni di trattamento acustico per la sala conferenze del Museo Egizio di Torino: (1) una soluzione con intonaco multistrato e (2) baffle parametrici in feltro. Per valutarne l'efficacia e generare le auralizzazioni in Ambisonics, è stato impiegato il software Ramsete. Un modello 3D fotogrammetrico ha permesso di realizzare i rendering 3D. Entrambi sono stati utilizzati nei test di intelligibilità del parlato condotti nell'Audio Space Lab del Politecnico di Torino, dotato di sistema Ambisonics e Oculus Meta Quest 2. Ai partecipanti è stato anche chiesto di esprimere la loro preferenza tra gli scenari visivi ottenuti con le due soluzioni di trattamento. I risultati hanno mostrato una preferenza per la seconda soluzione, nonostante non siano state riscontrate differenze acustiche significative.

Parole chiave: integrazione audiovisiva, auralizzazione, esperienza immersiva, intelligibilità del parlato, strumento di progettazione

1 | Introduction

Ensuring adequate Speech Intelligibility (SI) in environments designed for verbal communication, such as conference halls, is a crucial goal of acoustic design, as it directly influences both communication effectiveness and the overall listening experience [1]. This task becomes even more complex in existing spaces, particularly within historical buildings, where any acoustic intervention must carefully balance performance enhancement with the preservation of architectural value and heritage constraints. Conventional design methods are based on objective acoustic parameters and numerical simulations to estimate the impact of various interventions, providing a reliable but partial perspective [2]. While such methods are essential for evaluating acoustic performance, they frequently neglect the perceptual aspects of intelligibility, which are fundamental in actual listening situations. Recent advances in immersive tools, including virtual reality (VR) and auralisation, provide novel possibilities to enhance evaluation methods and to narrow the gap between numerical simulations and perceptual experience [3,4]. The present work investigates how these technologies can be incorporated into the design process through the comparison of two corrective solutions for the conference room of the Museo Egizio di Torino, with the dual objective of enhancing prediction reliability and supporting perceptual validation. Recent studies have investigated the integration of VR in architectural acoustics to improve design evaluation [5-9]. The possibility of previewing an architectural project together with an accurate reproduction of its expected acoustics allows users to experience the environment prior to construction, minimizing the need for design changes or post-construction repairs to address issues [5,6]. VR can be integrated into acoustic design through different strategies. A real-time approach offers flexibility by enabling the movement of sound sources, though with lower precision compared to pre-calculated wave-based simulations. Nonetheless, both methods contribute positively to the design process [7]. Such an approach necessarily involves collaboration across multiple disciplines, with each area of expertise contributing to achieve a faithful reproduction that enhances the realism and immersive quality of the virtual environment. Accurately modelling historical buildings is particularly challenging, as geometric detail plays a crucial role. Simplification techniques can be applied to reduce computational load while preserving an adequate level of visual fidelity [8]. Equally important is the accurate reproduction of the room's acoustic response. A robust strategy consists in calibrating the Geometrical Acoustic (GA) model on the basis of in-situ measurements, using them as a benchmark for the subsequent simulations [9].

In this framework, the article explores the role of VR techniques as a supporting tool in the acoustic design workflow. A central aspect of the study is to examine whether alternative visual configurations, under identical acoustic conditions, can affect user preferences.

A preliminary version of this research was presented at the International Symposium on Musical and Room Acoustics (ISMRA 2025). Compared to that contribution, the present work expands on several methodological aspects: the iterative calibration process is described in greater detail, the simulation parameters are explicitly reported to ensure reproducibility, and additional information is provided on the perceptual test protocol and statistical analysis.

2 | Methodology

The present study aims to integrate audio-visual (AV) VR techniques into architectural design using acoustic simulations with the GA Ramsete software. Third-Order Ambisonics (3OA) auralisations of two design proposals were realised and then compared in perceptual tests combining different AV conditions in five scenarios. Specifically, SI tests were carried out on normal hearing subjects. The experiment took place within the Audio Space Lab (ASL) of the Politecnico di Torino.

2.1 | The case study

The study focused on the conference room of the Museo Egizio di Torino. It is a room with essential furnishings for dissemination: a stage with a desk and projection screen at the back, rows of movable seats with two side benches, and a control desk positioned opposite the stage, used to manage the lighting system installed on the central elliptical structure. The measurements setup used within this room was in accordance with the Italian standard UNI 11532-2 [10] (Fig. 1).

These measures highlighted various criticalities affecting SI, a fundamental requirement in conference halls. In particular, the objective parameters T_{30} and C_{50} showed deviations from the optimal ranges established by current regulations (see Fig. 2).

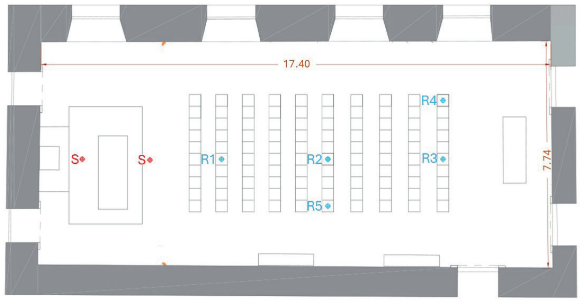

Fig. 1 – Set up measurement inside the conference room

Configurazione utilizzata nelle misurazioni della sala conferenze

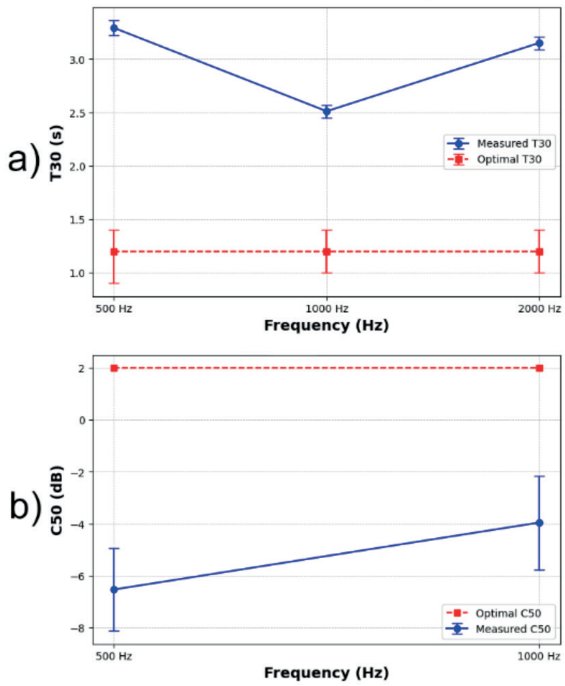

Fig. 2 – Measured vs. Optimal T30 (a) and C50 (b) within the conference room

Confronto tra i valori misurati e ottimali di T30 (a) e C50 (b) all'interno della sala conferenze

To address these shortcomings, an acoustic correction project was developed, resulting in two alternative proposals: one focused solely on acoustic treatment, while the other one introduced parametric baffles designed to improve SI and simultaneously enrich the visual quality of the space. The first design uses two types of fibrous sound absorbing panels covered with plaster arranged along three walls of the room (Fig. 3). In contrast, the second solution (Fig. 4) integrates fewer sound-absorbing panels with acoustic baffles, arranged parametrically in a horseshoe shape on three sides of the room. The two solutions are acoustically equivalent. Both are designed to improve SI by bringing objective parameters within the optimal ranges established by regulations.

Fig. 3 – First designed solution with plastered absorbing panels
Prima proposta progettuale con pannelli fonoassorbenti rivestiti d'intonaco

Fig. 4 – Second designed solution with parametric baffles
Seconda proposta progettuale parametrica con pannelli fonoas-sorbenti sospesi

2.2 | Simulations

The performance of the two corrective strategies was assessed through simulations carried out with the GA software Ramsete. This tool employs a pyramid-tracing approach, which reduces the risk of multiple-detection artefacts compared to traditional cone-tracing [11]. An additional capability of Ramsete is the generation of Ambisonics impulse responses (IRs), a key element for producing the auralisations employed in the perceptual tests. The software, however, is compatible exclusively with 3D models composed of triangular meshes. For this reason, the original model had to be adapted by restructuring the meshes into separate layers, each associated with the specific materials present in the room (e.g. plaster, floor, glass, etc.).

The Ramsete simulations were carried out following the guidelines provided by the software developer. The main parameters were: subdivision level = 10 (corresponding to 8×2^N pyramids per source, i.e. 81920 pyramids in this case), simulation time = 4.2 s (similar to Reverberation Time of the environment), time resolution = 0.001 s, and diffraction order = 2. Unlike other ray-tracing tools, Ramsete does not allow direct control of the number of rays, which is indirectly set by the subdivision level. These parameters ensured a sufficiently detailed representation of the sound field while maintaining reasonable computation times.

2.2.1 | Calibration

The calibration was carried out through an iterative process. Starting from literature values for absorption and scattering coefficients [12] (Tab. 1 and Tab. 2), a simulation was run and the resulting objective parameters (T_{20}, T_{30}, EDT, C_{50}, C_{80}, D_{50}, T_s) were compared against the corresponding in-situ measurements. Discrepancies guided the adjustment process (e.g., an overestimation of T_{30} at 1 kHz indicated the need to increase absorption in that band). The iteration was repeated until differences between simulated and measured values were consistently below the just noticeable difference (JND), resulting in average differences of 0.06 s for T_{30} and 1 dB for C_{50}. Only the absorption coefficients of the walls were modified (Tab. 3), since this material covers the largest surface of the hall and thus exerts the greatest influence on the overall absorption behaviour.

Tab. 1 – Absorption Coefficients from literature [12]
Coefficienti di assorbimento da letteratura

	125 Hz	*250 Hz*	*500 Hz*	*1000 Hz*	*2000 Hz*	*4000 Hz*	*8000 Hz*
Photo/canvas	0.42	0.21	0.10	0.08	0.06	0.06	0.06
Closet	0.14	0.12	0.10	0.08	0.06	0.06	0.03
Window	0.12	0.10	0.08	0.06	0.05	0.03	0.03
Wall	0.02	0.03	0.05	0.05	0.04	0.03	0.03
Black Wall	0.01	0.01	0.05	0.35	0.18	0.09	0.09
Floor	0.01	0.01	0.02	0.02	0.01	0.01	0.01
Door	0.11	0.08	0.06	0.06	0.05	0.03	0.02
Chair	0.22	0.24	0.22	0.20	0.18	0.14	0.12
Table	0.31	0.28	0.22	0.19	0.15	0.10	0.10
Stage	0.31	0.28	0.22	0.19	0.15	0.10	0.10
Projection sheet	0.14	0.12	0.10	0.08	0.06	0.06	0.03

Tab. 2 – Scattering Coefficients from literature [12]
Coefficienti di diffusione da letteratura

	125 Hz	*250 Hz*	*500 Hz*	*1000 Hz*	*2000 Hz*	*4000 Hz*	*8000 Hz*
Photo/canvas	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Closet	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Window	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Wall	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Black Wall	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Floor	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Door	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Chair	0.01	0.05	0.21	0.39	0.50	0.56	0.58
Table	0.01	0.05	0.21	0.39	0.50	0.56	0.58
Stage	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Projection sheet	0.02	0.03	0.06	0.15	0.26	0.32	0.36

Tab. 3 – Walls' absorption coefficients before and after calibration
Coefficienti di assorbimento delle pareti prima e dopo la calibrazione

	125 Hz	*250 Hz*	*500 Hz*	*1000 Hz*	*2000 Hz*	*4000 Hz*	*8000 Hz*
Before calibration	0.02	0.03	0.05	0.05	0.04	0.03	0.03
After calibration	0.02	0.04	0.05	0.06	0.06	0.08	0.12

2.2.2 | Comparison between the two designed proposals

The subsequent phase of the workflow consisted in virtually applying the two corrective solutions. For the first design, it was enough to modify the wall material properties and insert the absorption coefficients provided by manufacturer. The second design, instead, required integrating the parametric baffle system into the room model. This configuration envisaged the installation of 164 baffles, distributed with variable spacing between 10cm and 24 cm. Furthermore, each baffle is characterised by two different depths: when divided vertically into two sections, the upper section is twice as deep as the lower one (as can be seen in Fig. 5 [13]). Due to its geometric intricacy, this design would have generated an overly detailed model, rendering it impractical for processing within the GA software. To overcome this limitation, a simplified version of the model was developed, preserving the appropriate absorption and scattering coefficients so as to remain consistent with the original design intent. The simplified model was produced in Rhinoceros and was composed of a set of parallelepipeds, used to approximate the geometry of the panels. Additionally, due to the thickness difference between the upper and the lower sections, each part was modelled as a halved-thickness parallelepiped enclosing the lower sections of the baffles. These simplified geometrical elements were acoustically characterized using the absorption and scattering data obtained from reverberation chamber measurements on the panels' material [13], and such properties were subsequently implemented in the simulations (Tab. 4).

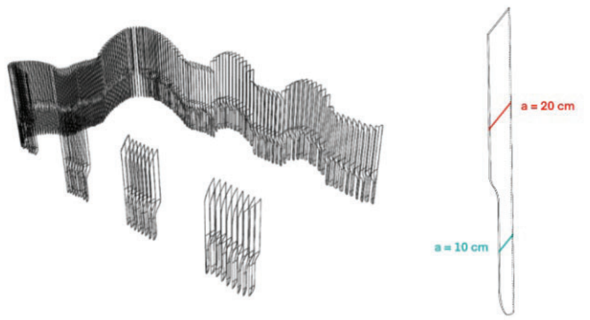

Fig. 5 – Parametric baffles structure
Struttura parametrica dei pannelli fonoassorbenti sospesi

Tab. 4: Baffles' Absorption (α) and Scattering (s) coefficients
Coefficienti di assorbimento e di diffusione dei pannelli fonoassorbenti sospesi

Hz	*63*	*125*	*250*	*500*	*1000*	*2000*	*4000*	*8000*
α	0.21	0.62	0.98	0.99	0.98	0.97	0.96	0.96
s	0.01	0.02	0.10	0.40	0.60	0.70	0.77	0.80

The simulation outcomes for the two redesigned configurations validated the effectiveness of both corrective interventions. Moreover, the results showed that the two solutions produced comparable effects, as illustrated in Fig. 6 for T_{30}.

2.2.3 | Visualisation

At the same time, the 3D environment was created. Using a photogrammetric model of the conference room [8], visual renderings were produced on Blender in an equirectangular format for visualization through a head-mounted display (HMD). Since the study also aimed to assess whether the visual component influenced the test results, two versions of the room were rendered: one with baffles and one without. Fig. 3 and Fig. 4 represent the previews, rendered with Blender, of the room in the two configurations.

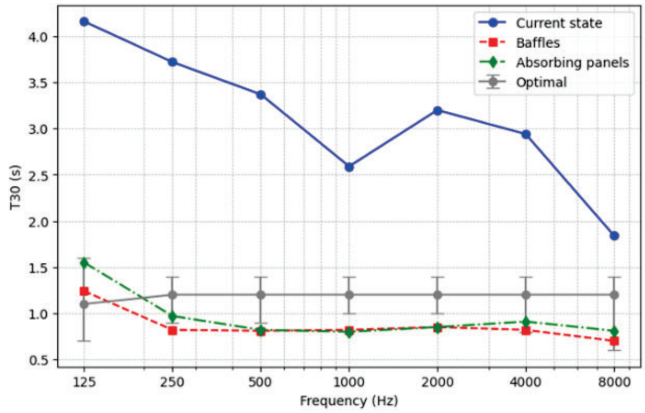

Fig. 6 – T_{30} - Comparison between the current state and the two correction strategies
T_{30} - Confronto tra lo stato di fatto e le due proposte progettuali

2.3 | Subjective Investigation

A perceptual study was carried out to assess how the proposed acoustic treatments were experienced by listeners. The decision to include visual cues was inspired by previous research [14,15], which demonstrated that both contextual and source-related visual information can influence sound localization [14] and acceptance of the auditory illusion [15].

In designing the experiment, five audiovisual scenarios were defined by combining two visual conditions (room without baffles and room with baffles) with three acoustic ones (untreated room and the two corrective interventions).

The combination of these AV conditions resulted in the creation of five scenarios used during the testing phase:

1. VNT + ANT - Visually Not Treated, Acoustically Not Treated.
2. VNT + ATP - Visually Not Treated, Acoustically Treated with Panels.
3. VNT + ATB - Visually Not Treated, Acoustically Treated with Baffles.
4. VT + ATP - Visually Treated, Acoustically Treated with Panels.
5. VT + ATB - Visually Treated, Acoustically Treated with Baffles.

This structure allowed us to examine not only the efficiency of the acoustic treatments but also the potential impact of visual immersion on SI.

2.3.1 | *Experimental Environment and Test Subjects*

Fig. 7 – A picture of the Audio Space Lab at Politecnico di Torino
Fotografia dell'Audio Space Lab al Politecnico di Torino

The experiment took place within the ASL of Politecnico di Torino, a compact listening environment treated for acoustics and designed for immersive reproduction (Fig. 7). The ASL complies with the ITU-R BS.1116-3 recommendation [16] and supports spatial playback up to 3OA. Synchronisation with an HMD enable the creation of a AV VR scenario. The reproduction system consists of 16 Genelec 8030B two-way active monitors, arranged on a spherical structure with 1.2-meters radius. The system is calibrated to achieve a sweet spot equalized in time, amplitude, and phase in the 90Hz-20kHz range. Loudspeakers are distributed across three rings, defined by the intersection of three parallel planes and a sphere:

- The central ring is at ear height and consists of eight equally spaced speakers, with the first one positioned directly in front of the listener.
- The upper and lower rings each contain four speakers, positioned at ±45° elevation relative to the central ring.

Additionally, the ASL features two Genelec 8531A three-way active monitors, positioned on the floor in the front, acting as subwoofers to extend the frequency range down to 30Hz. All loudspeakers are connected to a 32-channels sound card linked to a high-end desktop PC. Immersive video playback is managed through an Oculus Meta Quest 2. The AV synchronization is maintained with a MATLAB routing using Open-Source Control (OSC) protocol, integrating two different software: Bidule, and Unreal Engine.

The preliminary listening test involved ten normal hearing participants, six men and four women, aged 24-30 years. All participants were familiar with AV VR environments and HMDs, and 30% had previously participated in similar perceptual tests. Participants were recruited on a voluntary basis. Inclusion criteria required normal hearing and no uncorrected vision problems; participants who normally wore glasses were allowed to keep them during the test.

2.3.2 | *Speech Intelligibility Test*

Participants were positioned at the centre of the reproduction system and, while wearing an HMD, were exposed to each of the five AV scenarios. Their task consisted of repeating out loud the sentences uttered by the target source. In order to raise the level of difficulty, a competing noise source was added, which made the understanding of the sentences more demanding. The Signal-to-Noise Ratio (SNR) was set to -5dB, corresponding to a medium difficulty level.

Across all scenarios, the spatial arrangement was maintained unchanged, reflecting the layout commonly found in a conference room. The target speaker was located 4 meters away, placed behind the desk and oriented towards the listener. The interfering source was positioned 1.8 meters away from the listener, forming an angle of 120° relative to the frontal axis (as can be seen in Fig. 8). Moreover, no amplification system was employed during the experiment, and the speech material was reproduced at a "raised" vocal effort (70 dB SPL). For each scenario, participants listened to a set of 20 sentences randomly selected from the validated extended version of the Italian Matrix Sentence Test. The masking signal consisted of a phonetically balanced female voice, in line with common practice in SI assessments.

- The experiment was divided into three main phases:
- Training Phase – A progressively more challenging demo to familiarise participants with the test procedure.
- Testing Phase – Randomized exposure to the five AV scenarios, collecting SI scores (percentage of correctly understood words). Randomisation was applied at two levels: (i) the order of the audiovisual scenarios was changed for each subject, and (ii) the order of the sentences within each scenario was also varied. This approach ensured that neither learning effects nor presentation bias could systematically affect the results.
- Ranking Phase - Participants ranked their scenario preferences.

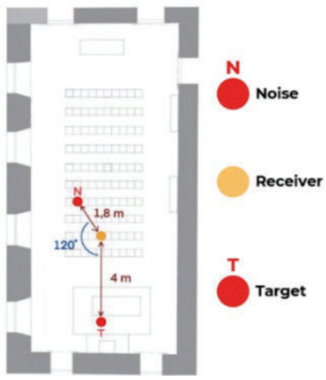

Fig. 8 – Spatial configuration used during the testing phase
Configurazione spaziale usata in fase di test

In summary, the adopted workflow combines acoustic simulation, auralisation, and 3D visual rendering to deliver an immersive and perceptually reliable assessment of alternative acoustic treatments. The workflow begins with the calibration of the RAMSETE acoustic model, carried out by assigning absorption and scattering coefficients from the literature and refining them through in-situ measurements performed in compliance with UNI 11532-2. The two design alternatives – parametric felt baffles and plaster-based acoustic panels – were assessed by producing 3OA IRs, later convolved with anechoic recordings to create spatialised auralisations. To accompany the auditory dimension with a faithful visual representation, a photogrammetric 3D model of the hall was processed and optimised in Blender, then imported into Unreal Engine for interactive real-time rendering. The virtual scene was configured to run on an Oculus Meta Quest 2 headset, enabling participants to explore the simulated acoustic environment. Furthermore, the auralisation chain was routed through the ASL playback system using Bidule, guaranteeing a precise reproduction of the spatial audio scene. MATLAB facilitates real-time interaction between the acoustic and visual domains through Open Sound Control (OSC) communication.

This methodology allows for an integrated evaluation of acoustic design options, merging the objective acoustic analyses with perceptual assessments within a controlled virtual framework. The combined use of numerical simulation, immersive auralisation and interactive 3D visualisation provides a robust framework for investigating the perceptual impact of different acoustic solutions.

3 | Results

3.1 | Objective Results

Tab. 5 – Speech Intelligibility Test Results
Risultati del test di intelligibilità del parlato

SCENARIOS		*VNT*			*VT*	
						
		ANT	*ATP*	*ATB*	*ATP*	*ATB*
Mean (%)	ALL	52.2	89.3	91.2	91.4	92.7
Std. Dev	ALL mean	11.38	3.83	6.12	3.44	3.65

In the Testing Phase, the number of words correctly recognised by each participant was collected, and the outcomes across scenarios are summarised in Tab. 5.

For each scenario, the mean number of correctly identified words per five-word sentence was computed. The findings revealed a substantial gain in intelligibility: in the untreated condition of the conference room, only 52.2% of the words were correctly perceived, whereas in the treated configurations comprehension rates rose to values between 89.3% and 92.7%. Because SI scores are bounded proportions, they were first converted into rationalized arcsine units (RAU) to stabilize variance and reduce ceiling effects. Statistical analysis was then conducted using a Linear Mixed Effects Model (LMM), with Scenario as a fixed effect and Subject as a random effect to account for repeated measures. Type III – F [17] tests were applied to assess the significance of Scenario, as they provide robust inference in the presence of unbalanced data and correlated observations. This approach is standard in perceptual studies where each listener is exposed to all conditions [18]. Fig. 9 presents the five scenarios ordered from the highest to the lowest intelligibility scores. Planned pairwise comparisons were then carried out between adjacent scenarios, under the directional hypothesis that intelligibility should not decrease when moving from worse to better conditions.

The following hypothesis was tested four times:

$$H_0: median(X1) \geq median(X2)$$

where X1 and X2 represent the scenario with the lower and higher intelligibility scores, respectively. For each comparison, the p-value was computed using a MATLAB routine (Tab. 6).

Tab. 6 – p-values for each hypothesis
Valori di p-value per ogni ipotesi

Hypothesis	*X1*	*X2*	*p-value*
1	VT + ATP	VT + ATB	0.17
2	VNT + ATB	VT + ATP	0.52
3	VNT + ATP	VNT + ATB	0.08
4	VNT + ANT	VNT + ATP	3.29×10^{-41}

The analysis of the calculated p-values indicates that only the last hypothesis (VNT+ANT vs. VNT+ATP) reached statistical significance, confirming that both acoustic treatments improved intelligibility compared to baseline. No significant differences emerged between the two designs, indicating that they are statistically equivalent within the variability of the sample.

3.2 | Subjective Results

During the Ranking Phase, the scenarios were ordered by the participants according to their preference, from the most to the last favoured. As reported in Tab. 7, the scenarios including baffles were generally preferred, while no appreciable subjective differences emerged between the two acoustic treatment strategies. A consistent outcome across all participants was the placement of the untreated room in the lowest rank. From an aesthetic standpoint, 90% of participants expressed a preference for the configuration with baffles, considering it as visually improving the environment. Additionally, the juxtaposition of the historic architecture with the contemporary aesthetic of the baffles was positively valued.

Tab. 7 – Ranking of the five tested scenarios by the participants
Classifica di gradimento dei cinque scenari testati dai partecipanti

SCENARIOS	*VNT*			*VT*	
					
	ANT	*ATP*	*ATB*	*ATP*	*ATB*
Overall	5th	3rd	3rd	1st	1st

4 | Discussion

The outcomes of this investigation support the feasibility of the two alternative corrective strategies, both capable of producing a marked improvement in SI. Objective simulations indicated that both interventions – panels alone and panels combined with felt baffles – yielded comparable results in terms of T_{30} and SI and complied with regulatory targets. Nevertheless, the perceptual evaluation highlighted a preference towards the baffle-based design, which was appreciated not only for its acoustic effectiveness but also for its visual qualities and the resulting sense of environmental comfort.

These results are consistent with the observations of Bradley [2], who underlined that conventional acoustic parameter should be complemented by subjective assessments to obtain a comprehensive evaluation of an environment. Indeed, the results indicate that while the two corrective strategies achieved equivalent measurable performance, the visual component – particularly the aesthetic contribution of the baffles – exerted a positive influence on the overall user experience. In this sense, the results also confirm the observations of Defays et al. [6] that visual information can modulate acoustic perception.

Furthermore, the study aligns with the work of Pind et al. [7], which demonstrated that VR can enhance the accuracy of acoustic predictions while integrating effectively within the overall design workflow. The investigation also underlines the value of combining diverse areas of expertise, leading to a workflow that merges acoustic simulation with visual design, in line with approaches proposed in the literature [8,9]. The adopted AV approach, integrating simulations calibrated on in-situ measurements with photogrammetry-based visual renderings, produced simulations that proved both realistic and reliable.

The results obtained in the ASL confirmed that both corrective strategies produced a significant enhancement of the SI score. A substantial contrast emerged between the untreated and the treated conditions, with intelligibility rising from 52.2% in the baseline scenario to 92.7% in the corrected configurations. The findings indicate that both corrective strategies were effective in enhancing speech intelligibility, with the statistical analysis confirming that their outcomes can be regarded as essentially equivalent.

It is noteworthy that a higher aesthetic evaluation was assigned exclusively to the baffle-based intervention. This outcome implies that beyond the acoustic enhancement, the visual aspect plays a crucial role in the overall perception of environmental quality. It should be underlined, however, that the intelligibility gains were not homogeneous among all participants. Such variability in performance may be attributed to individual factors, including prior familiarity with VR systems or previous exposure to similar perceptual tests. On average, the results point to an overall positive reception of the proposed corrective strategies.

It was also observed that the addition of visual stimuli affected the overall perception, with the visually enriched treatment contributing to a greater acceptance of the intervention. However, the role of visual factors in sound perception and overall acoustic behaviour in VR environments may require further study for a more in-depth understanding.

Although the outcomes of the simulations and laboratory trials were encouraging, the real-world applicability of the proposed solutions may be constrained by aspects such as cost, installation time, and visual impact. Baffled-based solutions might be more expensive and require a more complex design than acoustic panel solutions, which are likely to be easier to implement in existing environments. It is therefore important to consider the feasibility of these solutions within their architectural and practical context, particularly in a historic museum environment. In addition, the peculiar features of the museum – its geometry and furnishings in particular – played a decisive role in shaping the proposed acoustic solutions. The results obtained may not be directly transferable to other environments without further adaptation and specific analysis.

The findings indicate that numerical simulations benefit from being complemented with immersive evaluations, in order to obtain a more robust and reliable design process. Employing VR within the design workflow has the potential to streamline decision-making, minimise post-construction adjustments, and reduce overall project costs.

Future work foresees an expansion of the testing phase, with the introduction of new perceptual experiments designed to enable direct pairwise comparisons between alternative design configurations. As the present tests were carried out sequentially on an individual basis, it was not possible to clearly isolate the potential influence of the visual component on SI. Future studies should develop appropriate tests for use in acoustic design research.

5 | Conclusions

The investigation showed that embedding immersive VR technologies within the acoustic design workflow enables a more comprehensive assessment of corrective strategies in historic settings. Both corrective approaches were found to be effective in markedly enhancing SI, as confirmed by the outcomes of the acoustic simulations. Nonetheless, the perceptual evaluation highlighted a distinct aesthetic preference for the baffle-based solution, underlining the relevance of visual factors in shaping the overall perception of acoustic quality. These findings suggest that, although conventional acoustic parameters remain fundamental, their integration with immersive and experience-based assessments yields a more complete evaluation. The combined use of high-order Ambisonics auralisations and a photogrammetric 3D model provided a realistic reproduction of the conference hall, reducing the likelihood of costly post-construction modifications. Moreover, the multidisciplinary workflow adopted here illustrates how numerical simulation and subjective testing can complement each other, offering a more nuanced understanding of acoustic performance and user experience. Despite the limitations of the present study, the results provide a solid foundation for future research, particularly for direct pairwise comparisons aimed at better isolating the role of visual cues in auditory perception.

Overall, the proposed methodology represents a promising step towards a more reliable and user-centred acoustic design process in complex and heritage-sensitive environments, paving the way for further innovation in the field.

Conclusioni

L'indagine ha mostrato che l'integrazione delle tecnologie di realtà virtuale immersiva all'interno del flusso di lavoro di progettazione acustica consente una valutazione più completa delle strategie correttive in contesti storici. Entrambi gli approcci di intervento si sono rivelati efficaci nel migliorare in modo significativo l'intelligibilità del parlato, come confermato dai risultati delle simulazioni acustiche. Tuttavia, la valutazione percettiva ha evidenziato una chiara preferenza estetica per la soluzione con pannelli fonoassorbenti sospesi, sottolineando l'importanza dei fattori visivi nella percezione complessiva della qualità acustica. Tali evidenze suggeriscono che, sebbene i parametri acustici convenzionali rimangano fondamentali, la loro integrazione con valutazioni immersive ed esperienziali consente un'analisi più completa. L'impiego combinato di auralizzazioni Ambisonics di ordine elevato e di un modello fotogrammetrico 3D ha permesso una riproduzione realistica della sala conferenze, riducendo la probabilità di costose modifiche successive alla realizzazione. Inoltre, il flusso di lavoro multidisciplinare adottato dimostra come simulazioni numeriche e test soggettivi possano integrarsi in maniera complementare, offrendo una comprensione più approfondita sia delle prestazioni acustiche sia dell'esperienza dell'utente. Nonostante i limiti del presente studio, i risultati costituiscono una base solida per sviluppi futuri, in particolare per test di confronto diretto volti a isolare meglio il ruolo dei fattori visivi nella percezione uditiva.

In conclusione, la metodologia proposta rappresenta un passo promettente verso un processo di progettazione acustica più affidabile e orientato all'utente in ambienti complessi e sensibili dal punto di vista storico, aprendo la strada ad ulteriori innovazioni nel settore.

6 | Acknowledgement

The authors thank the Museo Egizio di Torino for the availability and support during the measurements in the case study and Filiberto Chiabrando, Lorenzo Teppati Losè (Department of Architecture and Design, Politecnico di Torino) and Stefano Rovera for their contribution to the construction of the model.

Bibliography

[1] R. Pisani, VALUTAZIONE DELLA QUALITA' ACUSTICA PER SALE ADIBITE ALL'ASCOLTO DEL PARLATO, (2010).

[2] J.S. Bradley, Review of objective room acoustics measures and future needs, Appl. Acoust. 72 (2011) 713–720. https://doi.org/10.1016/j.apacoust.2011.04.004.

[3] M. Ottley, Y. Chen, Use of Virtual Reality and auralization as a training tool for acoustic practitioners, in: Sydney, Australia, 2023: p. 015003. https://doi.org/10.1121/2.0001814.

[4] V. Hohmann, R. Paluch, M. Krueger, M. Meis, G. Grimm, The Virtual Reality Lab: Realization and Application of Virtual Sound Environments, Ear & Hearing 41 (2020) 31S-38S. https://doi.org/10.1097/AUD.0000000000000945.

[5] M. Azevedo, J. Sacks, Auralization as an Architectural Design Tool, (2014). https://doi.org/10.14279/depositonce-27.

[6] A. Defays, S. Safin, A. Billon, C. Decaestecker, N. Warzée, P. Leclercq, A.-S. Nyssen, Bimodal Interaction: The Role of Visual Information in Performing Acoustic Assessment in Architecture, TOERGJ 7 (2015) 13–20. https://doi.org/10.2174/1875934301407010013.

[7] F. Pind, C.-H. Jeong, H.S. Llopis, K. Kosikowski, J. Strømann-Andersen, Acoustic Virtual Reality – Methods and challenges, (2018).

[8] L. Shtrepi, S. Rovera, L.T. Lose', F. Chiabrando, A. Guastamacchia, A. Astolfi, ELABORAZIONE DI MODELLI 3D ACUSTICO-VISIVI PER LA PERCEZIONE DEL SUONO IN SPAZI STORICI, (2024).

[9] B.N.J. Postma, B.F.G. Katz, Creation and calibration method of acoustical models for historic virtual reality auralizations, Virtual Reality 19 (2015) 161–180. https://doi.org/10.1007/s10055-015-0275-3.

[10] UNI 11532-2, Caratteristiche acustiche interne di ambienti confinati - Metodi di progettazione e tecniche di valutazione - Parte 2: Settore scolastico (Internal acoustical characteristics of confined spaces - Design methods and evaluation techniques - Part 2: Educational sector), (2015).

[11] H. Hossam Eldien, The influence of an inclined line source close to building facades with balconies, Noise Cont Engng j 60 (2012) 363–373. https://doi.org/10.3397/1.3701016.

[12] T.J. Cox, P. D'Antonio, Acoustic absorbers and diffusers: theory, design, and application, 2. ed, Taylor & Francis, London, 2009.

[13] S. Rovera, L. Shtrepi, A. Astolfi, L.T. Lose', Application of the Performance-Based Design and Parametric Design Approach to the Acoustic Correction of Spaces Intended for Speech Listening: The Case Study of the Conference Hall at the Museo Egizio in Turin., Politecnico di Torino, 2024.

[14] A. Ahrens, K.D. Lund, M. Marschall, T. Dau, Sound source localization with varying amount of visual information in virtual reality, PLoS ONE 14 (2019) e0214603. https://doi.org/10.1371/journal.pone.0214603.

[15] A. Neidhardt, C. Schneiderwind, F. Klein, Perceptual Matching of Room Acoustics for Auditory Augmented Reality in Small Rooms - Literature Review and Theoretical Framework, Trends in Hearing 26 (2022) 23312165221092919. https://doi.org/10.1177/23312165221092919.

[16] ITU BS.1116-3, Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems, (2015).

[17] B.T. West, K.B. Welch, A.T. Gałecki, B.W. Gillespie, Linear Mixed Models: A Practical Guide Using Statistical Software, 3rd ed., Chapman and Hall/CRC, Boca Raton, 2022. https://doi.org/10.1201/9781003181064.

[18] A. Guastamacchia, F. Riente, L. Shtrepi, G.E. Puglisi, F. Pellerey, A. Astolfi, Speech intelligibility in reverberation based on audio-visual scenes recordings reproduced in a 3D virtual environment, Building and Environment 258 (2024) 111554. https://doi.org/10.1016/j.buildenv.2024.111554.