Audiovisual integration effect on the perception of the overall acoustic and architectural quality in a conference room

Proceedings of Meetings on Acoustics

Volume 58 http://acousticalsociety.org/

International Symposium on Musical and Room Acoustics

ISMRA 2025 NEW ORLEANS Playing the Room

Loyola University
New Orleans, Louisiana

24-28 May 2025

Architectural Acoustics

Audiovisual integration effect on the perception of the overall acoustic and architectural quality in a conference room

Louena Shtrepi, Davide Cetani and Stefano Rovera
Department of Energy, Politecnico di Torino, Torino, 10129, ITALY; louena.shtrepi@polito.it; davide.cetani@polito.it; stefano.rovera@studenti.polito.it

Lorenzo Teppati Losè and Filiberto Chiabrando
Department of Architecture and Design, Politecnico di Torino, Torino, 10129, ITALY; lorenzo.teppati@polito.it; filiberto.chiabrando@polito.it

Angela Guastamacchia, Marco Masoero and Arianna Astolfi
Department of Energy, Politecnico di Torino, Torino, 10129, ITALY; angela.guastamacchia@polito.it, marco.masoero@polito.it; arianna.astolfi@polito.it

Given the complexity of historical spaces, a detailed geometric 3D survey is essential to build an accurate visual model that realistically integrates the proposed design solutions for their acoustic enhancement. This work aims to investigate the audiovisual integration on the perception of the overall acoustic and architectural quality. It investigates two different design alternatives proposed for speech intelligibility improvements in a conference room inside the historical building at the Egyptian Museum of Turin. The first part of the research focused on acquiring and processing data for a detailed 3D geometric and acoustic model. Spatial room impulse responses have been measured in the existing conditions and have been used for the model calibration in the geometrical acoustic-based software Ramsete. Simulations have been performed for two solutions installations that share similar objective acoustic parameters: S1) plaster-based seamless acoustic system, and S2) parametrically shaped felt baffles. An immersive 3D render has been developed using a LiDAR and a photogrammetric-based model, that incorporate the designed acoustic elements. Listening tests have been presented through 3rd order ambisonics scenes at the Audio Space Lab (ASL) at Politecnico di Torino. The results show a significant preference for the baffle-based solution, despite no substantial differences in auditory perception.

Proceedings of Meetings on Acoustics, Vol. 58, 015007 (2025); https://doi.org/10.1121/2.0002143.

1. INTRODUCTION

Optimizing speech intelligibility (SI) in dedicated listening spaces, such as conference rooms, is a fundamental aspect of architectural design. It has a direct impact on communication clarity and listener experience.[1] This becomes even more challenging in existing environments, especially in historical buildings, where acoustic corrections must strike a balance between improving performance and respecting architectural and heritage constraints. Traditional approaches rely on objective acoustic parameters and simulation tools to predict the effectiveness of different treatments, providing a reliable but partial perspective. [2] Although these methods are essential for assessing acoustic performance, they often overlook the perceptual dimension of intelligibility, which plays a crucial role in real listening conditions. Recent innovations in immersive technologies, such as virtual reality (VR) and auralisation, offer new opportunities to refine the evaluation process and bridge the gap between numerical prediction and subjective experience. [3,4] This study explores the integration of these technologies into the acoustic design workflow by comparing two treatment solutions for a conference room at the Museo Egizio di Torino (Italy), with the aim of improving both prediction accuracy and perceptual validation.

Recent studies have investigated the integration of VR in architectural acoustics to improve design evaluation. [5-9] The ability to visualise an architectural design before its realization, combined with a faithful reproduction of the acoustics resulting from that design, gives the possibility of experiencing an environment before its realisation, minimizing the need for design changes or post-construction repairs to address issues. [5,6] There are several ways to implement VR in acoustic design. If flexibility is preferred, it is possible to use a real-time approach, allowing sound sources to be moved at the expense of the accuracy achievable with pre-calculated wave-based simulations. Both approaches add value to the building design process. [7] A multidisciplinary approach like this requires the collaboration of different sectors, each contributing their expertise to achieve a faithful reproduction that maximises the realism and immersion of the virtualised scenario. The quality of modelling historical buildings, where geometry becomes important, is not trivial. There are processes that can be used to simplify geometry to reduce computation time while still ensuring sufficient visual reproduction fidelity. [8] At the same time, it is essential that the acoustics of a real room are faithfully reproduced. An effective approach is to calibrate the Geometrical Acoustic (GA) model based on in-situ measurements and use the results as a reference for the simulation. [9] These tools have been shown to be of great importance and support when also aspects of dissemination towards other professionals and more general public engagement is necessary. [10] They have been further integrated in the acoustic and architectural design practice. [11,12]

In these terms, this article intends to investigate the use of VR techniques as a support within the acoustic design workflow. A key element is to investigate whether, under the same acoustic conditions, a different visual configuration can influence user preferences. The case study of a conference room was chosen given its spatial characteristics and cultural heritage relevance, i.e. it is an historical space with a vaulted ceiling in one of the most important buildings of the city of Turin. Traditional acoustic correction in these spaces is based on minimal interventions that do not vary significantly the aesthetics of the space. In this study, a more challenging acoustic correction has been proposed involving the installation of parametrically designed acoustic baffles, which do not only improve the acoustics of the room but are also considered as an enhancement of its aesthetic appearance. The above-mentioned technologies offer a more comprehensive framework to use in the comparison of the aesthetic-acoustic design alternatives and to report evidence-based data in the decision-making process that occurs when dealing with cultural heritage buildings.

2. METHODOLOGY

The present study aims to integrate audio-visual (AV) VR techniques into architectural design using acoustic simulations with the GA Ramsete software[13] of a conference hall (Figure 1). Third-Order Ambisonics (3OA) auralisations of two design proposals were realized 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.[14] The process of visual and acoustic model construction has been presented in detail from the 3D geometrical site survey based on photogrammetry and laser scanning methods to the visual and audio rendering obtained with Blender and Ramsete software.

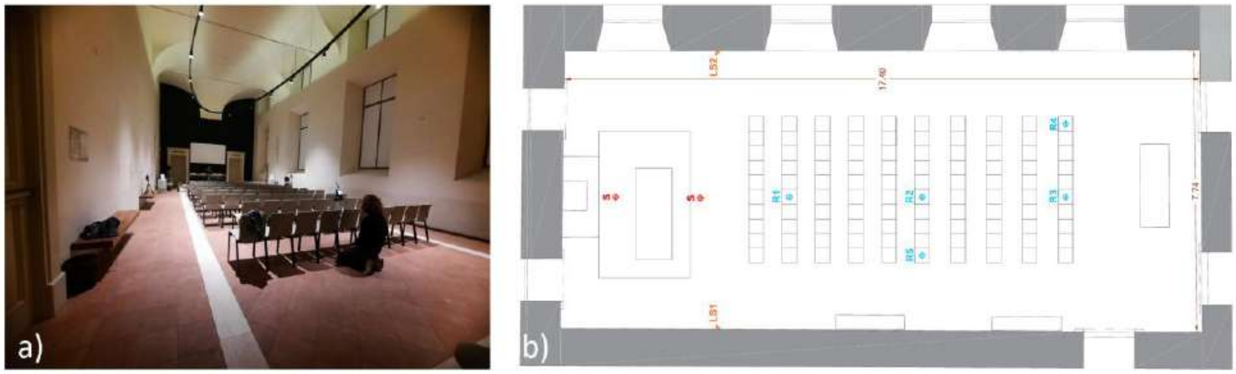

Figure 1: a) Photo of the conference room during the measurement campaign and b) measurement setup locations inside the conference room: sound source S and receivers R1-4.

A. CASE STUDY AND DESIGN PROPOSALS

The study focused on the conference hall of the Museo Egizio di Torino (Fig.1a and 1b). It is a room with essential furnishings for dissemination: a volume of $V=980\ m^3$, a stage with a desk, behind which there is a projection screen; movable seats (Nr=100) in the stalls and two side benches; and a desk on the opposite side of the stage, from which it is possible to control the lighting mounted on the elliptical structure suspended in the center of the room. In order to acoustically characterize the conference hall in unoccupied conditions, the reverberation time (T_{30}), properly averaged in space and frequency from 250 Hz to 4 kHz, was measured according to the ISO 3382-2:2008 standard. [15] The Brüel&Kjær 4292-L omnidirectional sound source driven by the Lab Gruppen LAB300 amplifier and the NTi Audio XL2 omnidirectional class-1 sound level meter were used for the measurements. The T_{30} resulted equal to 3.19 s ± 0.44 s, and C_{50} resulted equal to -5 dB ± 1.5 dB.

These measures revealed several issues that compromise speech intelligibility, which is crucial in conference room setting. The measured objective parameters, such as T_{30} and C_{50}, deviated from the optimal values defined by current regulations optimal values, i.e. on average T_{30} resulted 2 s longer than the optimal values and C_{50} resulted 5 dB lower than the optimal values for good speech comprehension in small conference halls according to the recent Italian standards on schools. [16]

To solve these discrepancies, an acoustic correction project was proposed, leading to the design of two different solutions: one that intervenes exclusively from an acoustic perspective, and another that involves the installation of parametrically designed acoustic baffles, which not only improve SI but also are considered as an enhancement of its aesthetic appearance. The first design uses two types of multilayer fibrous sound absorbing panels covered with acoustically transparent plaster and distributed along three walls of the room (detailed information is provided in part C section II).

B. VISUAL 3D MODEL

I. 3D MODEL ACQUISITION

The 3D model of the conference room (Figure 2) was created using an integrated process that combined photogrammetry and laser scanning to produce a detailed geometric model suitable for both acoustic simulation and immersive visual representation. [8,17] The approach has been used also in other studies[18] that highlight the process of simplification required to make use of the 3D model also in the acoustic simulations presented in the following sections. In the first phase, control points were obtained using topographic techniques, which were then used to georeference the photogrammetric and laser data. The photogrammetric survey (Figure 2a) was carried out using two cameras, Panasonic Lumix DMC-GH4 and Canon EOS 5DS R, and a drone, Parrot Anafi, and the acquired data was processed using Structure from Motion (SfM) within the Agisoft Metashape software.[17] Due to some difficulties related to the chromatic and tactile uniformity of some surfaces, a laser scanning campaign (Figure 2b and Figure 2c)) was integrated using a Faro Focus3D X330 instrument, with subsequent processing of the point clouds using FARO SCENE software. The data from the two methods were combined and converted into polygonal meshes, which were then optimised in Blender using the Remesh modifier, which made it possible to simplify the geometry while preserving the characteristics of the historic vault (Figure 2d). Finally, the Rhinoceros software was used to align the model with the reference system of the acoustic model so that it could be imported into the Ramsete simulation environment.

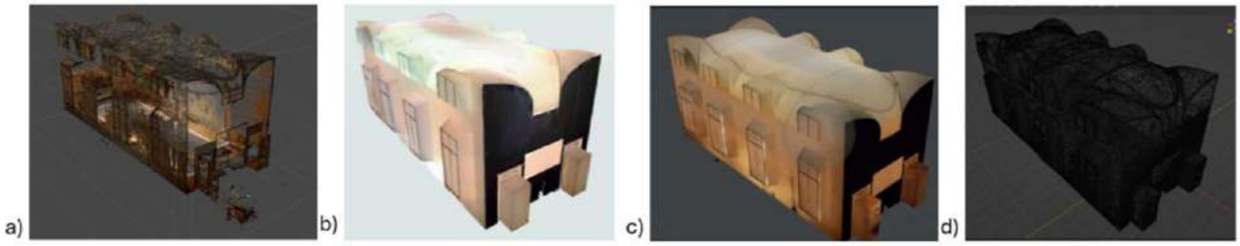

Figure 2: 3D model construction a) photogrammetric survey with dense point cloud representation obtained from image alignment, b) laser scanner with scan processing in FARO SCENE, c) 3D modeling with texture in Agisoft Metashape through the fusion of laser and photogrammetric data, d) import of the mesh model into Blender.

II. VISUALIZATION

Images of 360° in FARO SCENE have been used for visual rendering along with texture data collected with photogrammetry (Figure 3). 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 without (Figure 4a) and one with baffles (Figure 4b). In these renders, the setup used in the test phase was recreated and is described in detail below.

Figure 3: 360° image in FARO SCENE used for visual rendering along with texture data collected with photogrammetry.

C. ACOUSTIC SIMULATION

To predict the effectiveness of the two designs, the GA Ramsete software was used. Ramsete is based on pyramid-tracing algorithm, which is less prone to multiple detection issues compared to cone-tracing. [19] Another feature of Ramsete is the possibility to generate ambisonics Impulse Responses (IR), which are essential for auralisations performed during testing. However, this software only reads 3D models consisting of triangular meshes. Therefore, it was necessary to modify the 3D model to reorganize the meshes in different layers, corresponding to the materials present inside the room (plaster, floor, glass, etc.).

I. CALIBRATION

Table 1 shows acoustic properties taken from literature that were assigned to each material to create a GA model as a starting point. [20] Subsequently, the measurement setup was replicated in the simulation to compare the simulated objective parameters with the measured ones. This initiated an iterative calibration based on these parameters, ultimately aligning the model's acoustic behaviour with that of the conference room. This gave acceptable results with an average difference of 0.06s for the measured and calibrated T_{30}, while the measured and calibrated C_{50} differed by an average of 1 dB. During calibration, only the walls' absorption coefficients were modified, as can be seen in Table 2.

Table 1: Absorption (α) and Scattering (s) Coefficients from literature.[20]

Frequency [Hz]		125	250	500	1000	2000	4000	8000
Photo/canvas	α	0.42	0.21	0.10	0.08	0.06	0.06	0.06
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Closet	α	0.14	0.12	0.10	0.08	0.06	0.06	0.03
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Window	α	0.12	0.10	0.08	0.06	0.05	0.03	0.03
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Wall	α	0.02	0.03	0.05	0.05	0.04	0.03	0.03
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Black Wall	α	0.01	0.01	0.05	0.35	0.18	0.09	0.09
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Floor	α	0.01	0.01	0.02	0.02	0.01	0.01	0.01
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Door	α	0.11	0.08	0.06	0.06	0.05	0.03	0.02
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Chair	α	0.22	0.24	0.22	0.20	0.18	0.14	0.12
	s	0.01	0.05	0.21	0.39	0.50	0.56	0.58
Table	α	0.31	0.28	0.22	0.19	0.15	0.10	0.10
	s	0.01	0.05	0.21	0.39	0.50	0.56	0.58
Stage	α	0.31	0.28	0.22	0.19	0.15	0.10	0.10
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36
Projection sheet	α	0.14	0.12	0.10	0.08	0.06	0.06	0.03
	s	0.02	0.03	0.06	0.15	0.26	0.32	0.36

Table 2: Walls' absorption coefficients before and after calibration.

Frequency [Hz]	125	250	500	1000	2000	4000	8000
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

II. COMPARISON BETWEEN THE TWO DESIGNED PROPOSALS

The next step in the workflow involved the virtual implementation of the two correction strategies. In the first case, it was sufficient to change the material of the affected walls and input the supplier's absorption coefficients. In the second case, however, it was necessary to integrate the model of the baffles within the room. The project included the installation of 164 baffles with spacing ranging from 10 cm to 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. Figure 4c shows the baffle structure. The geometric complexity of this design would have resulted in an excessively detailed model, making it unmanageable for the GA software. To address this issue, a simplified model was created, maintaining the correct absorption and scattering coefficients to ensure fidelity to the original design. The simplified model, developed using Rhinoceros software, consists of a series of parallelepipeds representing 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. Table 3 shows baffles' material properties used in simulations. Baffles were characterized using measurements carried out in a reverberation room.[17]

Figure 4: Rendered images of the two solution strategies: a) with multilayer fibrous sound absorbing panels and b) parametrically designed acoustic baffles. c) Parametric baffles structure.

Table 3: Baffles' Absorption (α) and Scattering (s) coefficients.

Frequency [Hz]	125	250	500	1000	2000	4000	8000
α	0.62	0.98	0.99	0.98	0.97	0.96	0.96
s	0.02	0.10	0.40	0.60	0.70	0.77	0.80

Simulations of the two new hall configurations confirmed the effectiveness of both acoustic corrections. The results also indicate that the two strategies have a similar impact. Figure 5 illustrates this outcome for T_{30} (Figure 5a) and for C_{50} (Figure 5b).

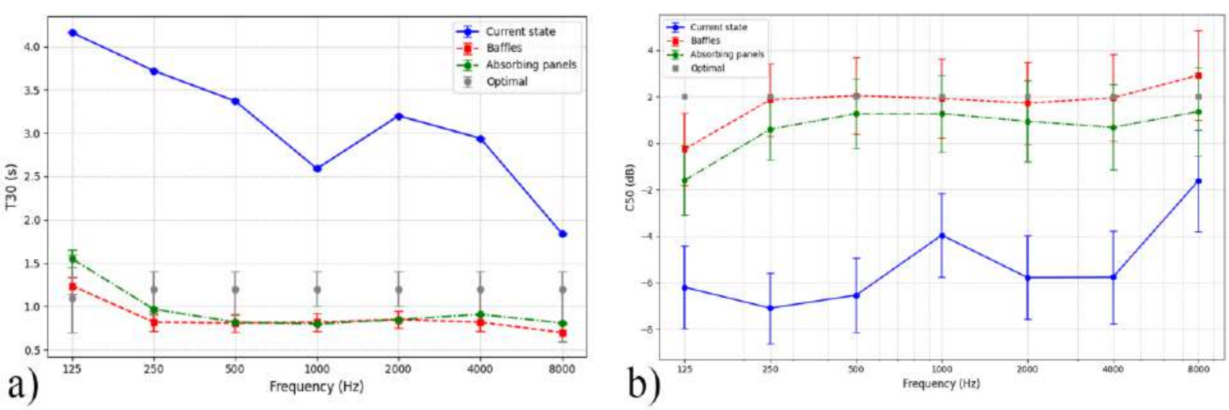

Figure 5: a) T_{30} - Comparison between the current state and the two correction strategies: absorbing panel and baffles; b) C_{50} - Comparison between the current state and the two correction strategies: absorbing panel and baffles.

D. SUBJECTIVE INVESTIGATION

A subjective investigation was conducted to evaluate the effectiveness of the acoustic interventions. The inclusion of visual stimuli was motivated by recent studies [21,22] suggesting that contextual and source-related visual cues influence sound localization[21] and acceptance of the auditory illusion[22].

The study identified different room configurations. Visually, the goal was to compare two setups: one with baffles and one representing the current state of the conference room. Acoustically, the goal was to compare three configurations: the current state of the room and the two proposed corrections. The combination of these AV conditions resulted in the creation of five scenarios used during the testing phase:

VNT + ANT - Visually Not Treated, Acoustically Not Treated.

VNT + ATP - Visually Not Treated, Acoustically Treated with Panels.

VNT + ATB - Visually Not Treated, Acoustically Treated with Baffles.
VT + ATP - Visually Treated, Acoustically Treated with Panels.
VT + ATB - Visually Treated, Acoustically Treated with Baffles.

This approach aimed to assess both the effectiveness of the two acoustic treatments and the influence of visual immersion on SI.

I. EXPERIMENTAL ENVIRONMENT AND TEST SUBJECTS

The experiment was carried out within the ASL[14], a small acoustically treated listening room at Politecnico di Torino. Figure 6a shows a picture of the laboratory. The ASL complies with the ITU-R BS.1116-3 recommendation[23] and is capable to reproduce immersive audio up to the 3OA. Additionally, the audio playback system can be synchronized with an HMD to create an immersive virtual 3D AV environment. The spherical playback array consists of 16 Genelec 8030B two-way active monitors, arranged in a 1.2-meters radius configuration. 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.

A pilot test was conducted with ten normal hearing participants, six men and four women, aged 24-30 years. All participants had prior experience with AV VR and HMDs, and 30% had previously participated in similar perceptual tests.

Figure 6: a) A picture of the Audio Space Lab at Politecnico di Torino and b) spatial configuration used during the testing phase.

II. DESIGN OF EXPERIMENT

During the test, participants were seated at the center of the listening setup, wearing an HMD, and experienced the five AV scenarios. They were required to repeat aloud sentences originated from the target source. To increase test difficulty, a noise source was introduced, making sentence comprehension more challenging. The Signal-to-Noise Ratio (SNR) was set to -5dB, corresponding to a medium difficulty level.

The spatial configuration was kept consistent across all scenarios, replicating a typical conference room setting. The target speaker was positioned behind the desk, facing the listener, at 4 meters. The noise source was placed 1.8 m from the listener, at a 120° angle from the frontal direction (Figure 6b). Additionally, the test was conducted without amplification, requiring a “raised” vocal effort (70 dB SPL). Each scenario was associated with a list of 20 random sentences, taken from the validated, extended version of the Italian Matrix Sentence Test. The female talker’s voice used for the noise source was phonetically balanced, as commonly adopted in speech recognition tests.

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).
- Ranking Phase - Participants ranked their scenario preferences.

To summarise the workflow, the proposed methodology integrates acoustic simulation, auralisation and 3D visual rendering to provide an immersive and perceptually accurate evaluation of different acoustic treatments. The process starts with the calibration of the acoustic model in RAMSETE, using both absorption and scattering coefficients derived from the literature and in-situ measurements according to the UNI 11532-2 standard. Two alternative solutions - felt baffles and acoustic panels - are evaluated by generating 3OA IRs, which are then convolved with anechoic signals to obtain spatialised auralisations. To complement the auditory experience with a realistic visual counterpart, a photogrammetric 3D model is developed and refined in Blender before being imported into Unreal Engine for real-time rendering. The virtual environment is designed for integration with Oculus Meta Quest 2, allowing for immersive exploration of the simulated acoustic spaces. In addition, the auralisation process is linked to the ASL playback system via Bidule, ensuring an accurate spatial audio experience. MATLAB facilitates real-time interaction between the acoustic and visual domains through Open Sound Control (OSC) communication.

Table 4: Speech Intelligibility Test Results.

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

This approach enables a comprehensive evaluation of acoustic treatments by combining objective acoustic analysis with subjective perceptual testing in a controlled virtual environment. 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.

III. STATISTICAL ANALYSIS

During the Testing Phase, the number of correctly understood words for each participant was recorded. For each scenario, the average number of correctly identified words per five-word sentence was calculated. To ensure statistical validity, a Type III – F test was conducted to assess the statistical significance. [24]

3. RESULTS

A. OBJECTIVE RESULTS

Table 4 presents a comparison across the results obtained in the different scenarios. The results indicate a remarkable improvement in intelligibility: in the scenario representing the current state of the conference room, participants correctly understood 52.2% of the words. In contrast, the acoustically treated scenarios showed an increase, with comprehension rates ranging from 89.3% to 92.7%. Figure 7 presents the five scenarios ordered from the highest to the lowest intelligibility scores. 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. Table 5 shows the calculated p-value for each hypothesis. The analysis of the calculated p-values indicates that only the last hypothesis can be rejected. This confirms that both noise correction designs are effective and that they are equivalent.

Table 5: p-values for each hypothesis.

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}

B. SUBJECTIVE RESULTS

In the Ranking Phase, participants were asked to rank the scenarios they experienced from most to least preferred. The results, shown in Table 6, indicate that the acoustically treated configurations with baffles were the most favoured; even from a subjective perspective, there are no significant differences between the two acoustic design solutions. The only common element is that the non-treated room was consistently ranked in last place. From an aesthetic perspective, 90% of participants preferred the version of the room with baffles, as they visually enhanced the space. Additionally, the contrast between the historic building and the modern design element of the baffles was highly appreciated.

Table 6: Ranking of the five tested scenarios by the participants.

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

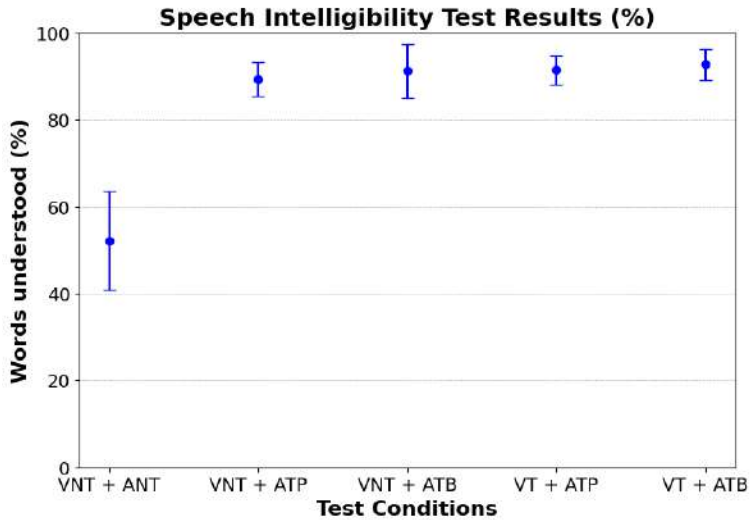

Figure 7: Percentage of understood words for each scenario during the testing phase.

4. DISCUSSION

The results of this study confirm the design feasibility of two equivalent acoustic correction treatments capable of significantly improving SI. The acoustic simulations showed that, from an objective point of view, both treatment – the one based solely on acoustic panels and the one integrated with felt baffles – have similar outcomes (in terms of T_{30} and SI test results) and are in line with the regulatory objectives. However, perceptual tests revealed a preference for the baffles solution; participants rated this configuration not only in terms of acoustic performance, but also for its aesthetic contribution and overall perception of environmental comfort.

These findings are in line with Bradley[2], who emphasises that standard acoustic parameters need to be supported in order to achieve a complete subjective assessment of the environment. Indeed, our data show that although the two interventions are equivalent in terms of measurable performance, the visual component – which, in the case of baffles, provides aesthetic enhancement – can positively influence the user experience. In this sense, the results also confirm the observations of Defays et al.[6] that visual information can modulate acoustic perception.

In addition, this research is in line with the work of Pind et al.[7], who have shown that the use of VR not only improves the accuracy of acoustic predictions but also fits seamlessly into the design workflow. The study also highlights the effectiveness of integrating different expertise, contributing to the creation of a workflow that combines acoustic simulation and visual design, as suggested by simulation methods in the literature. [8,9] The AV approach, which combines simulations based on in-situ measurements and visual renderings derived from photogrammetric models, has resulted in realistic and reliable simulations.

The ASL test has shown that both acoustic corrections have had a significant effect on the SI score. The difference between the untreated and acoustically treated configuration was remarkable, with the SI score improving from 52.2% in the untreated room to 92.7% in the acoustically treated configurations. This study suggests that both acoustic solutions are effective in improving SI, with essentially equivalent results between the two solutions, as the statistical test confirmed.

It is interesting to note that only the baffle treatment has received a higher aesthetic rating. This suggests that in addition to the acoustic improvement, the visual appearance plays a crucial role in the overall perception of environmental quality. However, it is important to note that the improvement in intelligibility was not uniform across all participants. Variations in scores may be due to individual differences such as familiarity with VR technology, or previous experience with this type of test. Nevertheless, the average scores indicate a generally positive response to the proposed solutions.

It should also be noted that the integration of the visual stimuli has had an impact on the overall perception, with the visually enhanced acoustic treatment also influencing on the 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.

While the simulations and laboratory tests have been promising, the practical applicability of the proposed solutions could be influenced by factors such as cost, implementation time and aesthetic 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 specific characteristics of the museum - such as the geometry and furnishings – have significantly influenced the acoustic solutions. The results obtained may not be directly transferable to other environments without further adaptation and specific analysis.

This study suggests that numerical simulations should be complemented by immersive evaluations to achieve a more reliable design process. The use of VR could improve the decision-making process, reduce the need for post-construction modifications and lower overall costs.

For future developments, the extension of the testing phase is planned. In particular, new perceptual tests will be conducted to allow direct pairwise comparisons between different design configurations. Since the tests in this study were conducted sequentially as individual tests, it was not possible to definitively determine whether the visual component had any effect on SI. Future studies should develop appropriate tests for use in acoustic design research.

5. CONCLUSION

This study demonstrated that the integration of immersive VR technologies into the acoustic design workflow can provide a more comprehensive evaluation of corrective interventions in historic environments. Both acoustic correction strategies proved effective in significantly improving SI, as confirmed by simulated objective parameters. However, perceptual tests revealed a clear aesthetic preference for the baffle solution, highlighting the importance of visual factors in the overall perception of acoustic quality. The results indicate that the traditional acoustic parameters remain indispensable but benefit greatly from the integration of immersive and experience-based evaluations. The combination of high-level Ambisonics auralisations with a calibrated, photogrammetric 3D model enabled a realistic reproduction of the museum's conference room, thus reducing the need for post-construction modifications and potentially lowering the overall project costs. Moreover, this multidisciplinary approach confirms that numerical simulations and subjective evaluations can complement each other, offering a more nuanced understanding of both acoustic performance and user experience. Although the study is subject to limitations, the results provide important highlights for future research. In particular, further studies involving direct pairwise comparisons are needed to better isolate the influence of visual cues on acoustic perception.

In conclusion, the proposed workflow represents a promising step towards more reliable and user-centred acoustic design in historic and complex environments, paving the way for further innovation and creating a common baseline in the field.

ACKNOWLEDGMENTS

The authors thank the Museo Egizio di Torino for the availability and support during the measurements in the case study and the listening test participants for their contribution to the research. They also thank Giuseppina Emma Puglisi, Elena Badino and Lorenzo Lavagna for their support during the acoustic measurements and Beatrice Tanduo for her support in the geometric survey.

REFERENCES

[1]Bradley JS., Sato H., Picard M. On the importance of early reflections for speech in rooms. Journal of the Acoustical Society of America, 2003, 113(6), 3233–3244.

[2]Bradley JS. Review of objective room acoustics measures and future needs. Appl Acoust 2011;72:713–20.

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

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

[5]Azevedo M, Sacks J. Auralization As An Architectural Design Tool, Proceedings of the EAA Joint Symposium on Auralization and Ambisonics 2014.

[6]Defays A, Safin S, Billon A, Decaestecker C, Warzée N, Leclercq P, et al. Bimodal Interaction: The Role of Visual Information in Performing Acoustic Assessment in Architecture. Ergon Open J 2015;7:13–20. https://doi.org/10.2174/1875934301407010013.

[7]Pind F, Jeong C-H, Llopis HS, Kosikowski K, Strømann-Andersen J. Acoustic Virtual Reality – Methods and challenges n.d.

[8]Shtrepi L, Rovera S, Lose' LT, Chiabrando F, Guastamacchia A, Astolfi A. Elaborazione di modelli 3d acustico-visivi per la percezione del suono in spazi storici. Proceedings of the 40th annual meeting of the Italian Acoustical Association, Taormina (Italy), June, 2024.

[9]Postma BNJ, Katz BFG. Creation and calibration method of acoustical models for historic virtual reality auralizations. Virtual Real 2015;19:161–80. https://doi.org/10.1007/s10055-015-0275-3.

[10]Katz B.F.G., Cros C., Peichert S., De Muynke J., The Past Has Ears at Notre-Dame: acoustic digital twins for research and narration, Digit Appl Archaeol Cult Herit, 34 (2024), pp. e00369:1-e00369:16, 10.1016/j.daach.2024.e00369

[11]Acentech, https://www.acentech.com/service/3dlistening/ (last (last access 20 April 2025).

[12]Arup, https://www.arup.com/services/digital-solutions-and-tools/soundlab/ (last (last access 20 April 2025).

[13]Farina A. Ramsete https://www.ramsete.com/ (last access 20 April 2025).

[14]Guastamacchia A, Rosso RG, Puglisi GE, Riente F, Shtrepi L, Astolfi A. Real and Virtual Lecture Rooms: Validation of a Virtual Reality System for the Perceptual Assessment of Room Acoustical Quality. Acoustics 2024;6:933–65. https://doi.org/10.3390/acoustics6040052.

[15]ISO 3382-2. Acoustics - Measurement of Room Acoustic Parameters – Part 2: Reverberation Time in Ordinary Rooms International Organization for Standardization, Genève (2008)

[16]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, 2020.

[17]Rovera S, 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. Master Thesis, Politecnico di Torino, 2024.

[18]Llorca-Bofí J, Heck J, and Vorlaender M, 3D pho-togrammetry for auralization - an approach to geometry simplification and material categorization, in Proceedings of BauSim Conference 2022, Weimar (Germany), Sept. 2022.

[19] Hossam Eldien H. The influence of an inclined line source close to building facades with balconies. Noise Control Eng J 2012;60:363–73. https://doi.org/10.3397/1.3701016.

[20] Cox TJ, D'Antonio P. Acoustic absorbers and diffusers: theory, design, and application. 2. ed. London: Taylor & Francis; 2009.

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

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

[23] International Telecommunication Union. ITU BS.1116-3, Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems. 2015.

[24] West BT, Welch KB, Gałecki AT, Gillespie BW. Linear Mixed Models: A Practical Guide Using Statistical Software. 3rd ed. Boca Raton: Chapman and Hall/CRC; 2022. https://doi.org/10.1201/9781003181064.